

On the hyperbolicity of Maxwell's equations with a local constitutive law

Volker Perlick*

*Physics Department, Lancaster University,
Lancaster LA1 4YB, UK*

and

The Cockcroft Institute, Warrington WA4 4AD, UK

Maxwell's equations are considered in metric-free form, with a local but otherwise arbitrary constitutive law. After splitting Maxwell's equations into evolution equations and constraints, we derive the characteristic equation and we discuss its properties in detail. We present several results that are relevant for the question of whether the evolution equations are hyperbolic, strongly hyperbolic or symmetric hyperbolic. In particular, we give a convenient characterisation of all constitutive laws for which the evolution equations are symmetric hyperbolic. The latter property is sufficient, but not necessary, for well-posedness of the initial-value problem. By way of example, we illustrate our results with the constitutive laws of biisotropic media and of Born-Infeld theory.

PACS numbers: 03.50.De 4.20.Cv

I. INTRODUCTION

The fact that Maxwell's equations can be formulated on a bare manifold that need not carry a metric or a connection was first observed by Kottler [1] and Cartan [2], and later also by van Dantzig [3] and Schroedinger [4], p. 24. It plays a central part in Post's [5] systematic study of the formal structure of electromagnetism. More recently, Hehl and Obukhov [6] have studied the metric-free (or pre-metric) approach to Maxwell's equations and its physical implications in great detail. It is the philosophy of Hehl and Obukhov to consider electromagnetism as more fundamental than gravity. In their approach, Maxwell's equations are formulated on a bare manifold. The constitutive law which connects the electromagnetic field strength with the electromagnetic excitation plays the role of a "space-time relation". In other words, the space-time geometry, which governs gravity, is coded in the constitutive law. By formulating a particular, very special, constitutive law one recovers a Lorentzian metric (at least up to a conformal factor) and, thus, the ordinary general-relativistic theory of gravity. It is then very natural to speculate that more general constitutive laws, which lead to more general geometric structures, could be considered as more general (hypothetical) theories of gravity. In particular, such generalised theories of gravity typically predict birefringence for light rays in vacuo (i.e., under the influence of gravity alone.)

The Hehl-Obukhov approach has some similarities, although more in philosophy than in mathematical technicality, with an idea of Newman and his collaborators (see, e.g. [7]) who suggest to view the equation of wave fronts as fundamental for gravity. A characteristic feature (and, maybe, a drawback) of both approaches is that the Lorentzian metric that is supposed to describe gravity can be fixed only up to a conformal factor, i.e., up to a strictly positive but otherwise undetermined scalar function.

The metric-free approach to Maxwell's equations gives a strong motivation for investigating which constitutive laws are physically reasonable and which ones are not. Here we want to discuss a criterion which is not mentioned in the book by Hehl and Obukhov: We want to characterise constitutive laws that yield a well-posed initial-value problem. A closely related property is the admittance of wavelike solutions, in a sense that is made precise in Section III below. Our results apply to the case that the constitutive law is interpreted as a spacetime relation (i.e., as the vacuum constitutive law in a generalised spacetime theory), but also to constitutive laws in a medium on a standard general-relativistic spacetime.

In contrast to Hehl and Obukhov, who restrict to local and linear constitutive laws throughout, we allow for nonlinear constitutive laws. However, we have to maintain the restriction to local constitutive laws which excludes, e.g., media with memory such as ferromagnets. The latter case would lead to integro-differential equations, for which an initial-value problem in the standard sense cannot be formulated, whereas a local constitutive law leads to first-order differential equations for appropriately chosen field components. In Section II we derive these differential equations and we decompose them into evolution equations and constraints. In Sections III and IV we derive and discuss the characteristic equation. The real roots of the characteristic equation determine the directions into which wavelike

*Author's address since November 2010: ZARM, Univ. Bremen, 28359 Bremen, Germany. Email: perlick@zarm.uni-bremen.de

solutions can travel and, thereby, the “light cones” of the theory. Non-real roots are associated with “evanescent modes”, i.e., with exponentially decaying solutions. In Section V we discuss the notions of hyperbolicity, strong hyperbolicity and symmetric hyperbolicity. Hyperbolicity requires that all roots of the characteristic equation are real, i.e., that evanescent modes do not occur. Strong hyperbolicity is a necessary and sufficient condition for the initial-value problem to be well-posed. Symmetric hyperbolicity is a sufficient but not a necessary condition for the initial-value problem to be well-posed. Having established these notions, we investigate some properties of the light cones for the case of hyperbolicity in Section VI. In Section VII we prove that the light cones are coordinate invariant which is not obvious from our derivation. The following three sections present some results that are useful for calculations: In Section VIII we discuss how the roots of the characteristic equation can actually be determined; in Section IX we demonstrate the invariance of the characteristic equation under certain changes of the constitutive law; and in Section X we derive an alternative form of the characteristic equation. The case that the light cones are invariant under temporal or spatial inversion is considered in Section XI, and the case that there is no birefringence is considered in Section XII. In Section XIII we characterise the class of all constitutive laws for which the evolution equations are symmetric hyperbolic. Finally, two examples are worked out in Section XIV: biisotropic media and Born-Infeld electrodynamics. In the conclusions we summarise the results that have been achieved so far, and we list some important questions that are still open.

II. MAXWELL’S EQUATIONS IN METRIC-FREE FORM

We consider a 4-dimensional bare manifold, with coordinates $x = (x^0, x^1, x^2, x^3)$. We use Einstein’s summation convention for latin indices running from 0 to 3 and for greek indices running from 1 to 3. We refer to x^0 as to the time coordinate and to x^μ as to the spatial coordinates. At present, this is just a convenient mode of expression. As we have no structure on our manifold, it does not make sense to ask whether the x^0 –lines are timelike or whether the hypersurfaces $x^0 = \text{constant}$ are spacelike. Later, however, we will discuss the question of whether initial values for the evolution part of Maxwell’s equations on the hypersurfaces $x^0 = \text{constant}$ determine a unique solution on an appropriate neighborhood. If this is true, one might view the covector dx^0 as “timelike”, in a sense determined by the evolution equations and not by a background structure. The covectors which are timelike in this sense turn out to form an open convex cone at each point of the manifold, see Section VI below.

In standard index notation, Maxwell’s equations read

$$\partial_{[a} F_{bc]}(x) = M_{abc}(x), \quad \partial_{[a} H_{bc]}(x) = J_{abc}(x), \quad (1)$$

where the square bracket denotes antisymmetrization. Here $F_{ab} = -F_{ba}$ is the electromagnetic field strength, $H_{ab} = -H_{ba}$ is the electromagnetic excitation, J_{abc} is the electric current and M_{abc} is a hypothetical magnetic current. (On physical grounds, there is good reason to assume that the latter is zero; however, we take it into account for the sake of generality.) Under coordinate transformations, these fields change according to

$$\tilde{F}_{ab} = \frac{\partial x^d}{\partial \tilde{x}^a} \frac{\partial x^e}{\partial \tilde{x}^b} F_{de}, \quad \tilde{M}_{abc} = \frac{\partial x^d}{\partial \tilde{x}^a} \frac{\partial x^e}{\partial \tilde{x}^b} \frac{\partial x^f}{\partial \tilde{x}^c} M_{def}, \quad (2)$$

$$\tilde{H}_{ab} = \frac{\det(\frac{\partial x}{\partial \tilde{x}})}{|\det(\frac{\partial x}{\partial \tilde{x}})|} \frac{\partial x^d}{\partial \tilde{x}^a} \frac{\partial x^e}{\partial \tilde{x}^b} H_{de}, \quad \tilde{J}_{abc} = \frac{\det(\frac{\partial x}{\partial \tilde{x}})}{|\det(\frac{\partial x}{\partial \tilde{x}})|} \frac{\partial x^d}{\partial \tilde{x}^a} \frac{\partial x^e}{\partial \tilde{x}^b} \frac{\partial x^f}{\partial \tilde{x}^c} J_{def}. \quad (3)$$

Thus, in the terminology of de Rham [8], F_{ab} and M_{abc} are even differential forms whereas H_{ab} and J_{abc} are odd differential forms. (An even differential form is the same as a totally antisymmetric covariant tensor field whereas an odd differential form is the same as a totally antisymmetric covariant pseudotensor field.)

We assume that the electric current J_{abc} and the magnetic current M_{abc} are given by equations of the form

$$J_{abc}(x) = j_{abc}(x, F(x), H(x)), \quad M_{abc}(x) = m_{abc}(x, F(x), H(x)). \quad (4)$$

Here it is essential that the values of the currents at x depend on x and on the values of field strength and excitation at x , but not on their derivatives. We will soon see that, under this assumption, the currents are irrelevant for the question of whether the initial value problem is well-posed. Therefore, they will play no role in our further discussion; we have allowed for non-zero currents only for the sake of generality. Note, however, that the functions j_{abc} and m_{abc} in (4) are not completely arbitrary. They must be consistent with the conservation laws

$$\partial_{[a} J_{bcd]}(x) = 0, \quad \partial_{[a} M_{bcd]}(x) = 0, \quad (5)$$

which are a consequence of (1).

In addition to (4) we assume that we have a constitutive law in the form of six scalar equations

$$\mathcal{F}_A(x, F(x), H(x)) = 0, \quad A = 1, \dots, 6 \quad (6)$$

which allow to express six of the twelve independent components $F_{ab}(x)$ and $H_{ab}(x)$ in terms of the remaining six. (We shall later specify the six components which are to be eliminated.) Again, it is essential that the constitutive law is local in the sense that knowledge of six components at a particular point x allows to express the remaining six components at this particular point x . In particular, it is essential that (6) does not involve derivatives of the field components.

We now separate the 8 equations (1) into two constraints

$$\partial_{[\mu} F_{\nu\sigma]}(x) = J_{\mu\nu\sigma}(x), \quad \partial_{[\mu} H_{\nu\sigma]}(x) = M_{\mu\nu\sigma}(x), \quad (7)$$

which do not contain any ∂_0 derivative, and six evolution equations

$$\partial_{[0} F_{\nu\sigma]}(x) = J_{0\nu\sigma}(x), \quad \partial_{[0} H_{\nu\sigma]}(x) = M_{0\nu\sigma}(x), \quad (8)$$

which do contain ∂_0 derivatives. For the well-posedness of the initial-value problem, only the evolution equations are relevant. The constraints restrict the allowed initial values. After solving the evolution equations with initial values that satisfy the constraints, one has to check whether the constraints are preserved. This is guaranteed if the currents (4) satisfy the conservation laws (5).

To link up with standard notation of electrodynamics, we decompose field strength and excitation in electric and magnetic parts.

$$\begin{aligned} F_{\nu 0} &= E_\nu, & F_{\mu\rho} &= \epsilon_{\mu\rho\sigma} B^\sigma, \\ H_{\nu 0} &= -\mathcal{H}_\nu, & H_{\mu\rho} &= \epsilon_{\mu\rho\sigma} D^\sigma, \end{aligned} \quad (9)$$

where $\epsilon_{\mu\rho\sigma}$ is the 3-dimensional Levi-Civita symbol, defined by the properties that it is totally antisymmetric and satisfies $\epsilon_{123} = 1$.

This puts the constraints (7) into the form

$$\partial_\rho D^\rho + \dots = 0, \quad \partial_\rho B^\rho + \dots = 0, \quad (10)$$

where the ellipses indicate terms that do not involve derivatives of the fields. The evolution equations (8) can be conveniently written in six-vector form. To that end, we write \vec{E} and $\vec{\mathcal{H}}$ for the three-column vectors with components E_1, E_2, E_3 and $\mathcal{H}_1, \mathcal{H}_2, \mathcal{H}_3$, respectively, and we write \vec{D} and \vec{B} for the three-column vectors with components D^1, D^2, D^3 and B^1, B^2, B^3 , respectively. Then the evolution equations (8) read

$$\partial_0 \begin{pmatrix} \vec{D} \\ \vec{B} \end{pmatrix} - \begin{pmatrix} \mathbf{0} & -\mathbf{A}^\rho \\ \mathbf{A}^\rho & \mathbf{0} \end{pmatrix} \partial_\rho \begin{pmatrix} \vec{E} \\ \vec{\mathcal{H}} \end{pmatrix} + \dots = 0. \quad (11)$$

Here the 3×3 matrices \mathbf{A}^ρ are defined by

$$\mathbf{A}^1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \quad \mathbf{A}^2 = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \mathbf{A}^3 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (12)$$

and, as in (10), the ellipses in (11) indicate terms that do not involve derivatives of the fields.

We shall now require that the constitutive equations (6) can be solved for $\vec{E}(x)$ and $\vec{\mathcal{H}}(x)$. If this is the case, we can eliminate \vec{E} and $\vec{\mathcal{H}}$ from (11) with the help of the constitutive equations, and we are left with a set of six first-order differential equations for the six dynamical variables \vec{B} and \vec{D} . If, on the other hand, the constitutive equations cannot be solved for $\vec{E}(x)$ and $\vec{\mathcal{H}}(x)$, the number of evolution equations does not coincide with the number of independent dynamical variables, so there is no chance to get a well-posed initial-value problem. We introduce the following terminology.

Definition 1. A coordinate system is called *admissible* if, in this coordinate system, the constitutive law (6) can be solved for $\vec{E}(x)$ and $\vec{\mathcal{H}}(x)$. A constitutive law is called *regular* at x if there is an admissible coordinate system on some neighborhood of x .

Henceforth we assume that we have a constitutive law that is regular at some chosen point, and we work in an admissible coordinate system defined on some open neighbourhood U of this point in M . We will see in Section VII that then, if U is chosen sufficiently small, almost all other coordinate systems on U are admissible as well, and we will investigate the behaviour under coordinate changes of all relevant quantities.

The assumption that (6) can be solved for $\vec{E}(x)$ and $\vec{\mathcal{H}}(x)$ results in equations of the form

$$\partial_\rho E_\mu(x) = \kappa_{\mu\tau}(x, \vec{D}(x), \vec{B}(x)) \partial_\rho D^\tau(x) + \chi_{\mu\tau}(x, \vec{D}(x), \vec{B}(x)) \partial_\rho B^\tau(x) + \dots \quad (13)$$

$$\partial_\rho \mathcal{H}_\mu(x) = \gamma_{\mu\tau}(x, \vec{D}(x), \vec{B}(x)) \partial_\rho D^\tau(x) + \nu_{\mu\tau}(x, \vec{D}(x), \vec{B}(x)) \partial_\rho B^\tau(x) + \dots \quad (14)$$

Here, as before, the ellipses indicate terms that do not involve derivatives of the fields. In the more particular case that the constitutive law is linear, the coefficients $\kappa_{\mu\tau}$, $\chi_{\mu\tau}$, $\gamma_{\mu\tau}$ and $\nu_{\mu\tau}$ depend only on x but not on the fields. In the following we denote by $\boldsymbol{\kappa}$, $\boldsymbol{\nu}$, $\boldsymbol{\chi}$ and $\boldsymbol{\gamma}$ the 3×3 matrices with components $\kappa_{\mu\tau}$, $\nu_{\mu\tau}$, $\chi_{\mu\tau}$ and $\gamma_{\mu\tau}$, respectively. $\boldsymbol{\kappa}$ is called the *impermeability* matrix, $\boldsymbol{\nu}$ is called the *impermeability* matrix, and $\boldsymbol{\chi}$ and $\boldsymbol{\gamma}$ are called the *magneto-electric cross-terms*. (Our notation follows Kong [9, 10].) The standard text-book formalism of electrodynamics is recovered if we assume that $\boldsymbol{\chi}$ and $\boldsymbol{\gamma}$ vanish and that $\boldsymbol{\kappa}$ and $\boldsymbol{\nu}$ depend only on x and are invertible. Then $\boldsymbol{\varepsilon} = \boldsymbol{\kappa}^{-1}$ is called the *permittivity* (or *dielectricity*) matrix and $\boldsymbol{\mu} = \boldsymbol{\nu}^{-1}$ is called the *permeability* matrix. A priori, however, there is no reason to assume that we can choose our coordinate system such that the magneto-electric cross-terms vanish and that $\boldsymbol{\kappa}$ and $\boldsymbol{\nu}$ are invertible. For a detailed discussion of media with magneto-electric cross-terms see O'Dell [11].

The four 3×3 matrices $\boldsymbol{\kappa}$, $\boldsymbol{\nu}$, $\boldsymbol{\chi}$ and $\boldsymbol{\gamma}$ can be combined into the 6×6 matrix

$$\mathbf{M} = \begin{pmatrix} \boldsymbol{\kappa} & \boldsymbol{\chi} \\ \boldsymbol{\gamma} & \boldsymbol{\nu} \end{pmatrix} \quad (15)$$

which we call the *constitutive matrix*.

With (13) and (14) inserted into (11), we get the following set of six evolution equations for the six dynamical variables \vec{D} and \vec{B} .

$$\partial_0 \begin{pmatrix} \vec{D} \\ \vec{B} \end{pmatrix} - \mathbf{L}^\rho \partial_\rho \begin{pmatrix} \vec{D} \\ \vec{B} \end{pmatrix} + \dots = 0, \quad (16)$$

where the 6×6 matrix

$$\mathbf{L}^\rho = \begin{pmatrix} \mathbf{0} & -\mathbf{A}^\rho \\ \mathbf{A}^\rho & \mathbf{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{\kappa} & \boldsymbol{\chi} \\ \boldsymbol{\gamma} & \boldsymbol{\nu} \end{pmatrix} \quad (17)$$

depends on x and, in the case of a non-linear constitutive law, also on $\vec{D}(x)$ and $\vec{B}(x)$. Thus, (16) is a quasilinear system of partial differential equations with non-constant coefficients.

Before we proceed further it is useful to add a remark on the fact that we had to solve the constitutive equations for \vec{E} and $\vec{\mathcal{H}}$, rather than for any other combination of field components, as a necessary condition for having a well-posed initial-value problem. It is sometimes argued (see, e.g., O'Dell [11], Section 2.1, or Hehl and Obukhov [12]) that one should solve the constitutive equations either for \vec{E} and \vec{B} , or for \vec{D} and $\vec{\mathcal{H}}$, because only then has the resulting equation a covariant (i.e., four-dimensional, coordinate-independent) meaning. According to this point of view, constitutive equations solved for other combinations of the field components are “a historical artifact” and “should be phased out from use” [12]. It is, indeed, true that the condition of solvability for \vec{E} and $\vec{\mathcal{H}}$ is not covariant. Nonetheless, it is precisely this condition which appears if we ask for a well-posed initial-value problem. This should not come as a surprise. The initial-value problem refers to a particular slicing of the spacetime into hypersurfaces $x^0 = \text{constant}$. It is *not* a problem that has a covariant answer; the initial-value problem is well-posed for some slicings, and not well-posed for others. So it is quite natural that non-covariant conditions play a role.

If the constitutive law can be solved not only for \vec{E} and $\vec{\mathcal{H}}$ but also for \vec{E} and \vec{B} , the impermeability matrix $\boldsymbol{\kappa}$ must be invertible. The constitutive matrix (15) can then be written in the form

$$\mathbf{M} = \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \boldsymbol{\gamma}\boldsymbol{\kappa}^{-1} & \mathbf{1} \end{pmatrix} \begin{pmatrix} \boldsymbol{\kappa} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\nu} - \boldsymbol{\gamma}\boldsymbol{\kappa}^{-1}\boldsymbol{\chi} \end{pmatrix} \begin{pmatrix} \mathbf{1} & \boldsymbol{\kappa}^{-1}\boldsymbol{\chi} \\ \mathbf{0} & \mathbf{1} \end{pmatrix} \quad (18)$$

as can be easily verified by multiplying out the right-hand side. (18) implies

$$\det(\mathbf{M}) = \det(\boldsymbol{\nu} - \boldsymbol{\gamma}\boldsymbol{\kappa}^{-1}\boldsymbol{\chi}) \det(\boldsymbol{\kappa}). \quad (19)$$

Analogously, if ν is invertible, we find

$$\mathbf{M} = \begin{pmatrix} \mathbf{1} & \chi\nu^{-1} \\ \mathbf{0} & \mathbf{1} \end{pmatrix} \begin{pmatrix} \kappa - \chi\nu^{-1}\gamma & \mathbf{0} \\ \mathbf{0} & \nu \end{pmatrix} \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \nu^{-1}\gamma & \mathbf{1} \end{pmatrix} \quad (20)$$

and hence

$$\det(\mathbf{M}) = \det(\kappa - \chi\nu^{-1}\gamma) \det(\nu). \quad (21)$$

Equations (18) and (20) are useful for calculating the inverse of \mathbf{M} .

III. APPROXIMATE-PLANE-WAVE SOLUTIONS OF MAXWELL'S EQUATIONS

Now we want to derive the characteristic equation of the evolution equations (16) and the resulting light cone structure. There are two quite different methods of how to do this. The first method, which goes back to Hadamard, investigates the directions in which discontinuities of the electromagnetic field can propagate. For linear constitutive laws on a bare manifold, this method is used in the book by Hehl and Obukhov [6]. The second method investigates the directions in which approximate-plane-wave solutions of Maxwell's equations can travel. This method was pioneered by Luneburg whose work is reviewed in the book by Kline and Kay [13]; their treatment is restricted to linear and isotropic constitutive laws on Minkowski spacetime.

Here we want to use the second method because it provides us with a clear physical interpretation of the characteristic equation. As there are no treatments in the literature that cover our situation – Maxwell's equations with a local but possibly nonlinear constitutive law on a bare manifold –, we give a detailed and self-contained exposition. Our first task is to define the notion of an “approximate-plane-wave solution”.

In standard electrodynamics on Minkowski spacetime, wave propagation can be studied in terms of plane harmonic waves. Maxwell's equations on a bare manifold do not admit plane-harmonic-wave solutions in general. However, they do admit such solutions in an approximative sense. To make this mathematically precise, we introduce the following terminology.

Definition 2. An *approximate-plane-wave family* with background field $\vec{D}(x)$, $\vec{B}(x)$ is a one-parameter family

$$\begin{pmatrix} \vec{\mathfrak{D}}(\alpha, x) \\ \vec{\mathfrak{B}}(\alpha, x) \end{pmatrix} = \begin{pmatrix} \vec{D}(x) \\ \vec{B}(x) \end{pmatrix} + \alpha \operatorname{Re} \left\{ \begin{pmatrix} \vec{d}(x) \\ \vec{b}(x) \end{pmatrix} + O(\alpha) \right\} \exp(iS(x)/\alpha) \quad (22)$$

with the following properties:

- (a) The coordinates $x = (x^0, x^1, x^2, x^3)$ range over an open neighborhood U of the manifold M and the parameter α ranges over the strictly positive real numbers, $\alpha \in \mathbb{R}^+$.
- (b) \vec{D} and \vec{B} are \mathbb{R}^3 valued C^∞ functions.
- (c) S is a real-valued C^∞ function whose gradient $dS(x) = \partial_a S(x) dx^a$ has no zeros on U . We refer to S as to the *eikonal function* of the approximate-plane-wave family.
- (d) \vec{d} and \vec{b} are \mathbb{C}^3 valued C^∞ functions with $(\vec{d}(x), \vec{b}(x)) \neq (\vec{0}, \vec{0})$ for all x in U ,

If $\partial_a S$, \vec{d} and \vec{b} are independent of x (in the chosen coordinate system) and the $O(\alpha)$ terms in (22) are zero, (22) is a background field with an α -dependent plane harmonic wave added; the wave covector of this plane harmonic wave is given by $k_a = \partial_a S/\alpha$. This observation gives the following interpretation to an arbitrary approximate-plane-wave family. On a sufficiently small neighborhood, $\partial_a S$, \vec{d} and \vec{b} differ arbitrarily little from constants, and for α sufficiently small the $O(\alpha)$ terms give arbitrarily small contributions. Thus, on a small neighborhood and for small α , (22) can be approximately viewed as a plane harmonic wave added to the background field. The smaller α , the more oscillations we have on the chosen neighborhood. We refer to $\alpha \rightarrow 0$ as to the *high-frequency limit*.

For the evaluation of approximate-plane-wave families the following simple lemma is crucial.

Lemma 3. Let S be the eikonal function of an approximate-plane-wave family and let u be a complex-valued continuous function defined on the same neighborhood U as S . Then $\lim_{\alpha \rightarrow 0} \operatorname{Re}\{u(x)e^{iS(x)/\alpha}\}$ exists for all x in U if and only if $u(x) = 0$ for all x in U .

Proof. If u is different from zero at some point in U , it is different from zero on an open subset V of U . For almost all x in V we have $S(x) \neq 0$, as dS has no zeros. As a consequence, for almost all x in V the limit does not exist. \square

We now consider an approximate-plane-wave family, given in coordinates $x = (x^0, x^1, x^2, x^3)$ that are admissible in the sense of Definition 1. We assume that the background fields $\vec{D}(x)$ and $\vec{B}(x)$ satisfy the constraints (10) and the evolution equations (16). We then say that the approximate-plane-wave family (22) is an N^{th} order *asymptotic solution* to the constraints (10) if

$$\lim_{\alpha \rightarrow 0} \left\{ \frac{1}{\alpha^N} (\partial_\rho \mathfrak{D}^\rho(\alpha, x) + \dots) \right\} = 0, \quad \lim_{\alpha \rightarrow 0} \left\{ \frac{1}{\alpha^N} (\partial_\rho \mathfrak{B}^\rho(\alpha, x) + \dots) \right\} = 0, \quad (23)$$

and to the evolution equations (16) if

$$\lim_{\alpha \rightarrow 0} \left\{ \frac{1}{\alpha^N} \left(\partial_0 \begin{pmatrix} \vec{\mathfrak{D}}(\alpha, x) \\ \vec{\mathfrak{B}}(\alpha, x) \end{pmatrix} - \mathbf{L}^\rho(x, \vec{\mathfrak{D}}(\alpha, x), \vec{\mathfrak{B}}(\alpha, x)) \partial_\rho \begin{pmatrix} \vec{\mathfrak{D}}(\alpha, x) \\ \vec{\mathfrak{B}}(\alpha, x) \end{pmatrix} + \dots \right) \right\} = \begin{pmatrix} \vec{0} \\ \vec{0} \end{pmatrix}. \quad (24)$$

Here the ellipses stand for the same terms as in (10) and (16), respectively. It is obvious that an N^{th} order asymptotic solution is automatically an M^{th} order asymptotic solution for all $M \leq N$. We want to investigate the lowest non-trivial order $N = 0$, which is known as the *geometric optics approximation*. (We will not consider asymptotic solutions of higher order, which allow to determine the $O(\alpha)$ terms in (22) iteratively.) Then in (23) and (24) the terms indicated by ellipses, which do not contain derivatives of the fields, give no contribution to the limit. Using Lemma 3, we find that an approximate-plane-wave family (22) is an asymptotic solution of order $N = 0$ to the constraints if and only if

$$\partial_\rho S(x) d^\rho(x) = 0, \quad \partial_\rho S(x) b^\rho(x) = 0, \quad (25)$$

and to the evolution equations if and only if

$$\partial_\rho S(x) \mathbf{L}^\rho(x, \vec{D}(x), \vec{B}(x)) \begin{pmatrix} \vec{d}(x) \\ \vec{b}(x) \end{pmatrix} = \partial_0 S(x) \begin{pmatrix} \vec{d}(x) \\ \vec{b}(x) \end{pmatrix}. \quad (26)$$

As the amplitudes are assumed to be non-zero, (26) means that $\partial_0 S(x)$ must be an eigenvalue of the 6×6 matrix $\partial_\rho S(x) \mathbf{L}^\rho(x, \vec{D}(x), \vec{B}(x))$, i.e.

$$\det \left(\partial_0 S(x) \mathbf{1} - \partial_\rho S(x) \mathbf{L}^\rho(x, \vec{D}(x), \vec{B}(x)) \right) = 0. \quad (27)$$

(27) is the *eikonal equation* of the evolution equations. It is a first order partial differential equation for S . Its coefficients depend not only on x but also on the background fields $\vec{D}(x)$ and $\vec{B}(x)$, unless we restrict to linear constitutive laws.

IV. THE CHARACTERISTIC EQUATION

Now we choose values for x , $\vec{D}(x)$ and $\vec{B}(x)$ and keep them fixed. We want to investigate the condition that there is a nonzero (\vec{d}, \vec{b}) such that (25) and (26) hold. This is an algebraic condition on the covector $dS = \partial_a S dx^a$, i.e., it determines a subset of the cotangent space at x . We will call any covector which satisfies this condition a “characteristic covector”. Hence, a covector is characteristic if and only if it is the gradient of an eikonal function of an approximate-plane-wave solution of order $N = 0$. The precise definition reads as follows. (As x , $\vec{D}(x)$, $\vec{B}(x)$ are now kept fixed, dependence on these quantities is no longer made explicit to ease notation.)

Definition 4. A covector $p_a dx^a$ is called *characteristic* if there is $(\vec{d}, \vec{b}) \neq (\vec{0}, \vec{0})$ in \mathbb{R}^6 such that

$$p_\rho d^\rho = 0, \quad p_\rho b^\rho = 0, \quad (28)$$

$$p_\rho \mathbf{L}^\rho \begin{pmatrix} \vec{d} \\ \vec{b} \end{pmatrix} = p_0 \begin{pmatrix} \vec{d} \\ \vec{b} \end{pmatrix}. \quad (29)$$

The set of all characteristic covectors is called the *characteristic variety*.

We will show in Section VII that this definition is coordinate-independent. Clearly, a necessary condition for $p_a dx^a$ to be characteristic is that it satisfies

$$\det(p_0 \mathbf{1} - p_\rho \mathbf{L}^\rho) = 0. \quad (30)$$

(30) is called the *characteristic equation* and its left-hand side is called the *characteristic polynomial* of the evolution equations. Note that (30) makes sure that (29) admits a non-trivial solution but does not take (28) into account.

By writing covectors as $p_a dx^a$ we have introduced canonical momentum coordinates p_a in each cotangent space which are conjugate to our admissible coordinates x^a . If x^0 can be interpreted as a temporal coordinate and the x^μ as spatial coordinates, we refer to p_0 as to the *frequency* and to \vec{p} as to the *spatial wave covector* (for dimensional reasons, one may put in a factor \hbar); here and in the following, \vec{p} stands for the three-column vector with components p_1, p_2, p_3 .

The left-hand side of (30) is a sixth order homogeneous polynomial in the variables p_a . For each $\vec{p} \in \mathbb{R}^3$, it is a sixth order polynomial in the variable p_0 . As such, it has six complex roots which will be denoted by $\omega_A(\vec{p})$ where $A = 1, \dots, 6$. In this notation, the characteristic equation (30) takes the form

$$\prod_{A=1}^6 (p_0 - \omega_A(\vec{p})) = 0. \quad (31)$$

The $\omega_A(\vec{p})$ are the eigenvalues of the matrix $p_\rho \mathbf{L}^\rho$,

$$p_\rho \mathbf{L}^\rho \begin{pmatrix} \vec{d}_A(\vec{p}) \\ \vec{b}_A(\vec{p}) \end{pmatrix} = \omega_A(\vec{p}) \begin{pmatrix} \vec{d}_A(\vec{p}) \\ \vec{b}_A(\vec{p}) \end{pmatrix}. \quad (32)$$

In general they are complex, and so are the components of the eigenvectors $\vec{d}_A(\vec{p})$ and $\vec{b}_A(\vec{p})$. Only real $\omega_A(\vec{p})$ are related to approximate-plane-wave solutions, because the eikonal function is supposed to be a real function. (Non-real eigenvalues are associated not with oscillating modes but rather with so-called *evanescent modes*, i.e., with exponentially decaying fields.) From the real solutions $\omega_A(\vec{p})$ of the eigenvalue problem (32) we have to single out those for which the eigenvectors satisfy the constraints (28); this will give us all the characteristic covectors in the form $\omega_A(\vec{p}) dx^0 + p_\mu dx^\mu$.

If \vec{p} runs over $\mathbb{R}^3 \setminus \{\vec{0}\}$, $p_\rho \mathbf{A}^\rho$ runs over all non-zero antisymmetric 3×3 matrices. Any such matrix has a one-dimensional kernel. As we can read from (17), this implies that the kernel of the matrix $p_\rho \mathbf{L}^\rho$ must be at least two-dimensional, hence $p_0 = 0$ is an eigenvalue of algebraic multiplicity ≥ 2 and the characteristic equation is of the form

$$p_0^2 \prod_{A=1}^4 (p_0 - \omega_A(\vec{p})) = 0. \quad (33)$$

The eigenvalues $\omega_A(\vec{p})$, for $A = 1, \dots, 4$, are in general complex and some of them may be zero as the algebraic multiplicity of the eigenvalue $p_0 = 0$ may be bigger than 2. If \mathbf{M} is invertible, the kernel of the matrix $p_\rho \mathbf{L}^\rho$ is precisely two-dimensional, hence $p_0 = 0$ is an eigenvalue of geometric multiplicity equal to 2; however, even in this case the algebraic multiplicity may be bigger than 2.

We will now show that, if $\omega_A(\vec{p}) \neq 0$, the corresponding eigenvectors automatically satisfy the constraints (28).

Proposition 1. If (32) holds with $\omega_A(\vec{p}) \neq 0$, the $\vec{d}_A(\vec{p})$ and $\vec{b}_A(\vec{p})$ satisfy $d_A^\rho(\vec{p}) p_\rho = b_A^\rho(\vec{p}) p_\rho = 0$.

Proof. The eigenvalue equation (32) implies that, for all complex numbers a and b ,

$$\begin{pmatrix} a \vec{p} \\ b \vec{p} \end{pmatrix} \cdot \begin{pmatrix} \mathbf{0} & -p_\rho \mathbf{A}^\rho \\ p_\rho \mathbf{A}^\rho & \mathbf{0} \end{pmatrix} \begin{pmatrix} \kappa & \chi \\ \gamma & \nu \end{pmatrix} \begin{pmatrix} \vec{d}_A(\vec{p}) \\ \vec{b}_A(\vec{p}) \end{pmatrix} = \begin{pmatrix} a \vec{p} \\ b \vec{p} \end{pmatrix} \cdot \omega_A(\vec{p}) \begin{pmatrix} \vec{d}_A(\vec{p}) \\ \vec{b}_A(\vec{p}) \end{pmatrix}, \quad (34)$$

where the dot denotes the standard scalar product in \mathbb{C}^6 . Using the antisymmetry of $p_\rho \mathbf{A}^\rho$, this can be rewritten as

$$\begin{pmatrix} \mathbf{0} & -p_\rho \mathbf{A}^\rho \\ p_\rho \mathbf{A}^\rho & \mathbf{0} \end{pmatrix} \begin{pmatrix} a \vec{p} \\ b \vec{p} \end{pmatrix} \cdot \begin{pmatrix} \kappa & \chi \\ \gamma & \nu \end{pmatrix} \begin{pmatrix} \vec{d}_A(\vec{p}) \\ \vec{b}_A(\vec{p}) \end{pmatrix} = \begin{pmatrix} a \vec{p} \\ b \vec{p} \end{pmatrix} \cdot \omega_A(\vec{p}) \begin{pmatrix} \vec{d}_A(\vec{p}) \\ \vec{b}_A(\vec{p}) \end{pmatrix}. \quad (35)$$

As \vec{p} spans the kernel of $p_\rho \mathbf{A}^\rho$, the left-hand side vanishes, so the right-hand side has to vanish for all complex numbers a and b . As $\omega_A(\vec{p})$ is non-zero, this completes the proof. \square

One may interpret the constraints (28) as saying that \vec{d} and \vec{b} must be *transverse*. By Proposition 1, this transversality condition follows already from the evolution equations for modes with non-zero frequency. It is interesting to note that such a transversality condition does not hold for \vec{e} and \vec{h} ; on a manifold without a metric, transversality cannot even be formulated for these fields because one would have to raise an index. This is another indication that the choice of \vec{D} and \vec{B} as the dynamical variables is the natural one, cf. Kong [9] and [10], Section 3.3, where an analogue of Proposition 1 is discussed for the special case of homogeneous linear media on Minkowski spacetime.

If the algebraic multiplicity of zero, as an eigenvalue of $p_\rho \mathbf{L}^\rho$, is equal to 2, Proposition 1 guarantees that a real covector $p_a dx^a$ is a characteristic covector if and only if it satisfies the *reduced characteristic equation*

$$\prod_{A=1}^4 (p_0 - \omega_A(\vec{p})) = \frac{1}{p_0^2} \det(p_0 \mathbf{1} - p_\rho \mathbf{L}^\rho) = 0. \quad (36)$$

Whereas (30) is the “characteristic equation of the evolution equations”, (36) can then be properly called the “characteristic equation of the full Maxwell equations” (evolution equations plus constraints).

The situation is very much more inconvenient if, for some $\vec{p} \in \mathbb{R}^3 \setminus \{\vec{0}\}$, the algebraic multiplicity of zero, as an eigenvalue of $p_\rho \mathbf{L}^\rho$, is > 2 . In this case one of the four roots $\omega_A(\vec{p})$ of the reduced characteristic equation (36) is zero. The corresponding covector $\omega_a(\vec{p})dx^0 + p_\mu dx^\mu$ may or may not be characteristic, i.e., the eigenvectors may or may not satisfy the constraints. If they do, we have an approximate-plane-wave solution of Maxwell’s equations of order $N = 0$ with zero frequency. One would conclude from this observation that x^0 cannot be interpreted as a temporal coordinate. Therefore it seems reasonable to discard such cases as unphysical, i.e., to restrict to cases where the coordinates can be chosen such that zero-frequency modes do not occur.

If κ is invertible, it is convenient to write the reduced characteristic polynomial (36) in the form

$$\prod_{A=1}^4 (p_0 - \omega_A(\vec{p})) = \det(\kappa) \mathcal{G}^{abcd} p_a p_b p_c p_d \quad (37)$$

so that the characteristic equation reads $\mathcal{G}^{abcd} p_a p_b p_c p_d = 0$. The reason for introducing the factor $\det(\kappa)$ is that the coefficients \mathcal{G}^{abcd} transform like a tensor density, as will be proven in Section VII. Note, however, that in the case of a non-linear constitutive law the \mathcal{G}^{abcd} depend not only on x but also on $\vec{D}(x)$ and $\vec{B}(x)$. Following Hehl and Obukhov [6] we call \mathcal{G}^{abcd} the *Tamm-Rubilar* tensor density. For an interesting representation of the Tamm-Rubilar tensor density, using the adjugate (or classical adjoint) of a matrix, see Itin [14].

Before discussing the (reduced) characteristic equation in more detail, we add a word on terminology. We have used the term “characteristic equation” which is the standard notation in texts on partial differential equation. In physics texts one finds the alternative term *dispersion relation*. Some authors, e.g. Hehl and Obukhov [6], use the term *Fresnel equation* as another alternative. Traditionally, the term “Fresnel equation” is used in crystal optics for an equation that determines the index of refraction (or, equivalently, the phase velocity) in dependence of the spatial direction, see e.g. Born and Wolf [15], Sect. 15.2.2. This equation is, indeed, equivalent to the characteristic equation.

V. HYPERBOLICITY OF THE EVOLUTION EQUATIONS

In the preceding section we have seen that the characteristic equation (30), together with the constraints, determines the directions $p_a dx^a$ in the cotangent space in which wavelike solutions can propagate. We will now discuss that the characteristic equation (30) also contains all information that is necessary to decide whether or not the evolution equations determine a well-posed initial-value problem.

Recall that we denote by $\omega_A(\vec{p})$, for $A = 1, \dots, 6$, the six eigenvalues of the matrix $p_\rho \mathbf{L}^\rho$ which depends on x , $\vec{D}(x)$ and $\vec{B}(x)$. Two of these eigenvalues are zero, $\omega_5 = \omega_6 = 0$, the other four are, in general, complex. We now recall some standard terminology from the theory of partial differential equations.

The evolution equations are called *hyperbolic* if all eigenvalues $\omega_A(\vec{p})$ are real, for all \vec{p} in \mathbb{R}^3 . (This is the case if and only if, in the terminology of Gårding [16], the characteristic polynomial $\det(p_0 \mathbf{1} - p_\rho \mathbf{L}^\rho)$ is hyperbolic with respect to the covector dx^0 .) The evolution equations are called *strongly hyperbolic* if for each \vec{p} there is an invertible matrix $\mathbf{S}(\vec{p})$ such that $\mathbf{S}(\vec{p})^{-1} p_\rho \mathbf{L}^\rho \mathbf{S}(\vec{p})$ is symmetric. They are called *symmetric hyperbolic* if this is true with an \mathbf{S} that is independent of \vec{p} . Obviously the following implications hold: symmetric hyperbolic \Rightarrow strongly hyperbolic \Rightarrow hyperbolic.

Of course, the answer to the question of whether any of the three properties – hyperbolic, strongly hyperbolic or symmetric hyperbolic – holds, may vary in dependence of x , $\vec{D}(x)$ and $\vec{B}(x)$.

Hyperbolicity guarantees the unrestricted existence of approximate-plane-wave solutions. However, it is too weak to guarantee well-posedness of the initial-value problem for the class of differential equations – quasilinear with non-constant coefficients – to which our evolution equations belong.

The latter requires strong hyperbolicity. More precisely, if the evolution equations are strongly hyperbolic at some x , $\vec{D}(x)$ and $\vec{B}(x)$, the following holds true. Data for \vec{D} and \vec{B} on the hypersurface $x^0 = \text{constant}$ that take the prescribed values $\vec{D}(x)$ and $\vec{B}(x)$ at x determine a unique solution to the evolution equations on some neighborhood of x , and the solution depends on the data continuously. The data must be of Sobolev class H^s , for some $s \geq 3$, and continuity is meant with respect to the Sobolev H^s norm. For details and proofs the reader is referred to Taylor [17], Theorem 5.2D. (Note that Taylor uses the term “symmetrizable” instead of “strongly hyperbolic”.)

In the more special case of symmetric hyperbolicity we have not only continuous dependence of the solution on the data but in addition we can control the growth of the solution in terms of *energy inequalities*. Usually symmetric hyperbolicity is easier to check than hyperbolicity or strong hyperbolicity. In Section XIII below we will give a convenient characterisation of all constitutive laws that give symmetric hyperbolic evolution equations. For hyperbolic or strongly hyperbolic evolution equations, no such characterisation is known so far.

Lindell, Sihvola and Suchy [18], in an otherwise very useful article, claim that the eigenvalues $\omega_A(\vec{p})$ are real whenever the constitutive matrix \mathbf{M} is symmetric. This would give a very convenient sufficient condition for hyperbolicity. Unfortunately, the claim is wrong; a counter-example is

$$\mathbf{M} = \begin{pmatrix} \mathbf{0} & \mathbf{1} \\ \mathbf{1} & \mathbf{0} \end{pmatrix} \quad (38)$$

for which the matrix $p_\rho \mathbf{L}^\rho$ has eigenvectors $\omega_1(\vec{p}) = \omega_2(\vec{p}) = i|\vec{p}|$, $\omega_3(\vec{p}) = \omega_4(\vec{p}) = -i|\vec{p}|$ and $\omega_5(\vec{p}) = \omega_6(\vec{p}) = 0$. The error comes in eq. (65) of [18] where the authors divide by a quantity without paying attention to the fact that this quantity may be zero. (Note that Lindell, Sihvola and Suchy allow the constitutive matrix to be complex, in contrast to the formalism used here, and that their \mathbf{M} is our \mathbf{M}^{-1} . In the above argument we specified their reasoning to the case that \mathbf{M} is real and we used the obvious fact that an invertible real matrix is symmetric if and only if its inverse is symmetric.)

VI. THE LIGHT CONES

The homogeneity of the characteristic polynomial implies that the functions ω_A are positively homogeneous of degree one,

$$\omega_A(s\vec{p}) = s\omega_A(\vec{p}) \quad \text{for all } s > 0. \quad (39)$$

We now assume hyperbolicity, and we order the 4 roots of the reduced characteristic polynomial according to

$$\omega_1(\vec{p}) \geq \omega_2(\vec{p}) \geq \omega_3(\vec{p}) \geq \omega_4(\vec{p}). \quad (40)$$

This guarantees that the ω_A are continuous, but not necessarily smooth, functions. Thus, the reduced characteristic equation defines four connected sets

$$\mathcal{C}_A = \{ p_a dx^a \mid p_0 = \omega_A(\vec{p}) \}, \quad A = 1, 2, 3, 4, \quad (41)$$

in the (real) cotangent space of the chosen point x . Without our assumption of hyperbolicity one would have to consider the \mathcal{C}_A as subsets of the complexified cotangent space.

By (39), each \mathcal{C}_A is a cone in the sense that it is invariant under multiplication with positive real numbers. We refer to the \mathcal{C}_A as to the *four branches of the characteristic variety* or, shorter, as to the four *light cones*. By differentiating the characteristic equation we find immediately that the differential of $p_0 - \omega_A(\vec{p})$ is non-zero at any point of \mathcal{C}_A where \mathcal{C}_A does not meet one of the other light cones. Thus, \mathcal{C}_A is a 3-dimensional manifold at any such point. At an intersection point with some other light cone \mathcal{C}_B , however, \mathcal{C}_A need not be smooth. For instance, \mathcal{C}_A and \mathcal{C}_B may form a “conical singularity”, with pointed tips meeting head-on. This gives rise to the phenomenon of “conical refraction” whose observability has been a matter of vivid debate in the history of anisotropic optics; for a detailed discussion see e.g. Born and Wolf [15], p. 813–818.

Our ordering (40) implies that \mathcal{C}_4 is the image of \mathcal{C}_1 and \mathcal{C}_3 is the image of \mathcal{C}_2 under reflection at the origin in the cotangent space. This follows immediately from the fact that the characteristic polynomial is homogeneous. Thus, $\mathcal{C}_1 \cup \mathcal{C}_4$ is a double-cone in the sense that it is generated by straight lines through the origin, and so is $\mathcal{C}_2 \cup \mathcal{C}_3$. Note, however, that in general there is no reflection symmetry with respect to the plane $p_0 = 0$.

We will now show that the four light cones divide up into two past cones and two future cones, provided that the reduced characteristic polynomial has no zero roots. In the terminology explained in Section IV, the latter condition means that we prohibit zero-frequency modes.

Proposition 2. Assume that the four roots of the reduced characteristic polynomial are real and non-zero for all $\vec{p} \neq \vec{0}$, and that we order them according to (40). Then

$$\omega_1(\vec{p}) \geq \omega_2(\vec{p}) > 0 > \omega_3(\vec{p}) \geq \omega_4(\vec{p}) \quad (42)$$

holds for all $\vec{p} \neq \vec{0}$.

Proof. Assume that three of the four roots ω_A are positive at some \vec{p} . Owing to the homogeneity of the characteristic polynomial, three roots must be negative at $-\vec{p}$. However, as the ω_A are continuous on $\mathbb{R}^3 \setminus \{\vec{0}\}$, this is possible only if some ω_A has a zero somewhere on $\mathbb{R}^3 \setminus \{\vec{0}\}$ which contradicts our assumption. We have thus proven that it is impossible that three roots are positive. By the same token, it is impossible that three roots are negative, so we must have two positive and two negative roots, i.e., (42) must be true. \square

From this proposition we find the following Corollary.

Proposition 3. Assume that the evolution equations are strongly hyperbolic and that the constitutive matrix \mathbf{M} is invertible. Then (42) holds for all $\vec{p} \neq \vec{0}$.

Proof. What we have to prove is that our assumption prohibits zero-frequency modes, i.e., that $p_0 = 0$ has algebraic multiplicity equal to 2 as an eigenvalue of $p_\rho \mathbf{L}^\rho$ for all $\vec{p} \neq \vec{0}$. We have already observed in Section IV that invertibility of \mathbf{M} guarantees that the geometric multiplicity of $p_0 = 0$ is equal to 2. Strong hyperbolicity makes sure that algebraic and geometric multiplicity coincide. \square

Whenever we have hyperbolicity, \mathcal{C}_1 and \mathcal{C}_4 are the boundaries of *convex* open cones

$$\mathcal{Z}_1 = \{ p_a dx^a \mid p_0 > \omega_1(\vec{p}) \}, \quad \mathcal{Z}_4 = \{ p_a dx^a \mid p_0 < \omega_4(\vec{p}) \}. \quad (43)$$

The convexity of \mathcal{Z}_1 and \mathcal{Z}_4 is a general feature, following from hyperbolicity, as was already proven in Gårding's pioneering paper [16]. By contrast, \mathcal{C}_2 and \mathcal{C}_3 are not in general the boundaries of convex sets. One may call covectors in \mathcal{Z}_1 “future-pointing timelike” and covectors in \mathcal{Z}_4 “past-pointing timelike”. Of course, future and past interchange their roles under reflection of the coordinate x^0 . The fact that the light cones, as subsets of the cotangent space, are independent of the chosen coordinate system will be proven in Section VII.

VII. COORDINATE TRANSFORMATIONS

We have worked in a chosen admissible coordinate system throughout, and we will now investigate to what extent the results found are invariant with respect to coordinate transformations. In particular, we will verify that the light cones are invariant (i.e., coordinate-independent).

As the hypersurfaces $x^0 = \text{constant}$ play a distinguished role, it is useful to distinguish coordinate transformations that leave these hypersurfaces invariant. The most general such transformation induces on each cotangent space a linear transformation of the form

$$d\tilde{x}^\mu = a^\mu{}_\rho (dx^\rho + v^\rho dx^0), \quad d\tilde{x}^0 = c dx^0 \quad (44)$$

and on each tangent space the dual linear transformation

$$\frac{\partial}{\partial \tilde{x}^\mu} = b_\mu{}^\rho \frac{\partial}{\partial x^\rho}, \quad \frac{\partial}{\partial \tilde{x}^0} = \frac{1}{c} \left(\frac{\partial}{\partial x^0} - v^\rho \frac{\partial}{\partial x^\rho} \right). \quad (45)$$

Here $\mathbf{a} = (a^\mu{}_\rho)$ is an invertible 3×3 matrix, $\mathbf{b} = (b_\mu{}^\rho)$ is the transpose of its inverse (i.e. $\mathbf{a} \mathbf{b}^T = \mathbf{b}^T \mathbf{a} = \mathbf{1}$ or, in index notation, $a^\mu{}_\rho b_\nu{}^\rho = b_\nu{}^\rho a^\rho{}_\mu = \delta_\nu^\mu$), (v^1, v^2, v^3) is a real 3-tuple and c is a non-zero real number. We call (44) and (45) a *generalised Galilean transformation*. It reduces to a standard Galilean transformation if $(a^\mu{}_\rho)$ is orthogonal and $|c| = 1$. An arbitrary coordinate transformation induces on the cotangent and tangent spaces linear transformations that can be written as a generalised Galilean transformation (44) and (45) followed by a transformation of the form

$$d\hat{x}^\mu = d\tilde{x}^\mu, \quad d\hat{x}^0 = d\tilde{x}^0 + u_\sigma d\tilde{x}^\sigma, \quad (46)$$

$$\frac{\partial}{\partial \hat{x}^\mu} = \frac{\partial}{\partial \tilde{x}^\mu} - u_\mu \frac{\partial}{\partial \tilde{x}^0}, \quad \frac{\partial}{\partial \hat{x}^0} = \frac{\partial}{\partial \tilde{x}^0}, \quad (47)$$

where (u_1, u_2, u_3) is an arbitrary real 3-tuple.

From the transformation behaviour (2) and (3) of field strength and excitation we can calculate the transformation behaviour of the fields \vec{E} , \vec{B} , $\vec{\mathcal{H}}$ and \vec{D} as defined in (9), and thereupon of the constitutive matrix whose components are defined by (13) and (14). With respect to generalised Galilean transformations (44) and (45), we find

$$\begin{aligned} \tilde{E}_\nu &= \frac{1}{c} b_\nu{}^\tau (E_\tau + v^\rho \epsilon_{\rho\tau\sigma} B^\sigma), & \tilde{B}^\mu &= (\det(\mathbf{a}))^{-1} a^\mu{}_\nu B^\nu, \\ \tilde{\mathcal{H}}_\nu &= \frac{\det(\mathbf{a})}{|c| |\det(\mathbf{a})|} b_\nu{}^\tau (\mathcal{H}_\tau - v^\rho \epsilon_{\rho\tau\sigma} D^\sigma), & \tilde{D}^\mu &= \frac{c}{|c| |\det(\mathbf{a})|} a^\mu{}_\nu D^\nu, \end{aligned} \quad (48)$$

which yields, after some elementary algebra, the following transformation rule for the constitutive matrix.

$$\begin{pmatrix} \tilde{\kappa} & \tilde{\chi} \\ \tilde{\gamma} & \tilde{\nu} \end{pmatrix} = \frac{\det(\mathbf{a})}{c} \begin{pmatrix} \frac{c \det(\mathbf{a})}{|c \det(\mathbf{a})|} \mathbf{b} \kappa \mathbf{b}^T & \mathbf{b} (\chi + v^\sigma \mathbf{A}_\sigma) \mathbf{b}^T \\ \mathbf{b} (\gamma - v^\sigma \mathbf{A}_\sigma) \mathbf{b}^T & \frac{c \det(\mathbf{a})}{|c \det(\mathbf{a})|} \mathbf{b} \nu \mathbf{b}^T \end{pmatrix}. \quad (49)$$

Here the antisymmetric matrices $\mathbf{A}_\rho = \mathbf{A}^\rho$ are defined by (12). By (49), $\tilde{\mathbf{M}}$ is well-defined whenever \mathbf{M} is. Hence, a generalised Galilean transformation transforms admissible coordinate systems into admissible coordinate systems.

With these results at hand, and with the transformation of the canonical momentum coordinates

$$\tilde{p}_\mu = b_\mu{}^\rho p_\rho, \quad \tilde{p}_0 = \frac{1}{c} (p_0 - v^\rho p_\rho), \quad (50)$$

we can now calculate the transformation behaviour of the eigenvalues and eigenvectors of $p_\rho \mathbf{L}^\rho$.

Proposition 4. Under a generalised Galilean transformation (44) and (45), the eigenvalues and eigenvectors (32) of $p_\rho \mathbf{L}^\rho$ transform according to

$$\tilde{\omega}_A(\vec{p}) = \frac{1}{c} (\omega_A(\vec{p}) - v^\rho p_\rho), \quad (51)$$

$$\begin{pmatrix} \vec{d}_A(\vec{p}) \\ \vec{b}_A(\vec{p}) \end{pmatrix} = \det(\mathbf{a})^{-1} \begin{pmatrix} \mathbf{a} & \mathbf{0} \\ \mathbf{0} & \mathbf{a} \end{pmatrix} \begin{pmatrix} \frac{c \det(\mathbf{a})}{|c \det(\mathbf{a})|} \vec{d}_A(\vec{p}) \\ \vec{b}_A(\vec{p}) \end{pmatrix} \quad (52)$$

for $A = 1, 2, 3, 4$.

Proof. The transformation behaviour (49) of the constitutive matrix, together with (50), allows us to calculate the transformed matrix $\tilde{p}_\rho \tilde{\mathbf{L}}^\rho$,

$$\begin{pmatrix} \mathbf{a} & \mathbf{0} \\ \mathbf{0} & \mathbf{a} \end{pmatrix}^{-1} \tilde{p}_\rho \tilde{\mathbf{L}}^\rho \begin{pmatrix} \mathbf{a} & \mathbf{0} \\ \mathbf{0} & \mathbf{a} \end{pmatrix} = \frac{1}{c} p_\rho \begin{pmatrix} \mathbf{0} & -\mathbf{A}^\rho \\ \mathbf{A}^\rho & \mathbf{0} \end{pmatrix} \begin{pmatrix} \frac{c \det(\mathbf{a})}{|c \det(\mathbf{a})|} \kappa & \chi + v^\sigma \mathbf{A}_\sigma \\ \gamma - v^\sigma \mathbf{A}_\sigma & \frac{c \det(\mathbf{a})}{|c \det(\mathbf{a})|} \nu \end{pmatrix}. \quad (53)$$

Now assume that (32) is true. This eigenvalue equation can be equivalently rewritten as

$$p_\rho \begin{pmatrix} \mathbf{0} & -\mathbf{A}^\rho \\ \mathbf{A}^\rho & \mathbf{0} \end{pmatrix} \begin{pmatrix} \frac{c \det(\mathbf{a})}{|c \det(\mathbf{a})|} \kappa & \chi \\ \gamma & \frac{c \det(\mathbf{a})}{|c \det(\mathbf{a})|} \nu \end{pmatrix} \begin{pmatrix} \frac{c \det(\mathbf{a})}{|c \det(\mathbf{a})|} \vec{d}_A(\vec{p}) \\ \vec{b}_A(\vec{p}) \end{pmatrix} = \omega_A(\vec{p}) \begin{pmatrix} \frac{c \det(\mathbf{a})}{|c \det(\mathbf{a})|} \vec{d}_A(\vec{p}) \\ \vec{b}_A(\vec{p}) \end{pmatrix}. \quad (54)$$

With $\tilde{\omega}_A(\vec{p})$, $\vec{d}_A(\vec{p})$ and $\vec{b}_A(\vec{p})$ introduced by (51) and (52), we find from (53) and (54), with the help of Proposition 1, that (32) holds with all terms twiddled. \square

From this proposition we read that

$$\tilde{p}_\rho d\tilde{x}^\rho + \tilde{\omega}_A(\vec{p}) d\tilde{x}^0 = p_\rho dx^\rho + \omega_A(\vec{p}) dx^0 \quad (55)$$

which demonstrates that the light cones are invariant. The reduced characteristic polynomial transforms as

$$\prod_{A=1}^4 (\tilde{p}_0 - \tilde{\omega}_A(\vec{\tilde{p}})) = \frac{1}{c^4} \prod_{A=1}^4 (p_0 - \omega_A(\vec{p})) . \quad (56)$$

With

$$\det(\tilde{\kappa}) = \frac{|\det(\mathbf{a})|}{|c|} \det(\kappa) , \quad (57)$$

which follows from (49), this implies

$$\frac{1}{\det(\tilde{\kappa})} \prod_{A=1}^4 (\tilde{p}_0 - \tilde{\omega}_A(\vec{\tilde{p}})) = \frac{1}{|c \det(\mathbf{a})|} \frac{1}{\det(\kappa)} \prod_{A=1}^4 (p_0 - \omega_A(\vec{p})) . \quad (58)$$

This shows that under generalised Galilean transformations the \mathcal{G}^{abcd} , introduced in (37), transform as a tensor density of weight 1,

$$\tilde{\mathcal{G}}^{abcd} \tilde{p}_a \tilde{p}_b \tilde{p}_c \tilde{p}_d = \left| \det\left(\frac{\partial x}{\partial \tilde{x}}\right) \right| \mathcal{G}^{abcd} p_a p_b p_c p_d . \quad (59)$$

We now turn to transformations of the form (46) and (47). The fields change according to

$$\begin{aligned} \hat{E}_\nu &= \tilde{E}_\nu , & \hat{B}^\mu &= \tilde{B}^\mu + \epsilon^{\mu\nu\sigma} u_\nu \tilde{E}_\sigma , \\ \hat{\mathcal{H}}_\nu &= \tilde{\mathcal{H}}_\nu , & \hat{D}^\mu &= \tilde{D}^\mu - \epsilon^{\mu\nu\sigma} u_\nu \tilde{\mathcal{H}}_\sigma , \end{aligned} \quad (60)$$

where $\epsilon^{\mu\nu\sigma}$ is the contravariant Levi-Civita symbol, defined by the properties that it is totally antisymmetric and satisfies $\epsilon^{123} = 1$. From this we find the following transformation behaviour of the constitutive matrix.

$$\begin{pmatrix} \hat{\kappa} & \hat{\chi} \\ \hat{\gamma} & \hat{\nu} \end{pmatrix} = \begin{pmatrix} \tilde{\kappa} & \tilde{\chi} \\ \tilde{\gamma} & \tilde{\nu} \end{pmatrix} \left(\begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{pmatrix} - u_\rho \begin{pmatrix} \mathbf{0} & -\mathbf{A}^\rho \\ \mathbf{A}^\rho & \mathbf{0} \end{pmatrix} \begin{pmatrix} \tilde{\kappa} & \tilde{\chi} \\ \tilde{\gamma} & \tilde{\nu} \end{pmatrix} \right)^{-1} . \quad (61)$$

Thus, a coordinate transformation of the form (46) and (47) maps admissible coordinates into admissible coordinates if and only if the inverse matrix on the right-hand side of (61) exists. This is true for almost all values of (u_1, u_2, u_3) , namely whenever $d\tilde{x}^0 + u_\rho d\tilde{x}^\rho$ is non-characteristic.

From (61) and the transformation behaviour of the momentum coordinates,

$$\hat{p}_\mu = \tilde{p}_\mu - u_\mu \tilde{p}_0 , \quad \hat{p}_0 = \tilde{p}_0 , \quad (62)$$

we find the transformation behaviour of the characteristic matrix,

$$\hat{p}_0 \mathbf{1} - \hat{p}_\rho \hat{\mathbf{L}}^\rho = (\tilde{p}_0 \mathbf{1} - \tilde{p}_\rho \tilde{\mathbf{L}}^\rho) (\mathbf{1} - u_\sigma \tilde{\mathbf{L}}^\sigma)^{-1} . \quad (63)$$

This demonstrates that the characteristic equations $\det(\hat{p}_0 \mathbf{1} - \hat{p}_\rho \hat{\mathbf{L}}^\rho) = 0$ and $\det(\tilde{p}_0 \mathbf{1} - \tilde{p}_\rho \tilde{\mathbf{L}}^\rho) = 0$ are equivalent, i.e., that the light cones are invariant with respect to coordinate transformations of the form (46) and (47). To verify the transformation behaviour of the Tamm-Rubilar tensor density, we assume that $\tilde{\kappa}$ and $\tilde{\mathbf{M}}$ are invertible. Then (61) takes the form

$$\hat{\mathbf{M}}^{-1} = \tilde{\mathbf{M}}^{-1} - u_\rho \begin{pmatrix} \mathbf{0} & -\mathbf{A}^\rho \\ \mathbf{A}^\rho & \mathbf{0} \end{pmatrix} . \quad (64)$$

After calculating the inverse matrices $\tilde{\mathbf{M}}^{-1}$ and $\hat{\mathbf{M}}^{-1}$ with the help of (18), the lower right-hand block of (64) yields

$$\hat{\nu} - \hat{\gamma} \hat{\kappa}^{-1} \hat{\chi} = \tilde{\nu} - \tilde{\gamma} \tilde{\kappa}^{-1} \tilde{\chi} , \quad (65)$$

which, by (19), implies

$$\det(\hat{\mathbf{M}}) \det(\tilde{\kappa}) = \det(\tilde{\mathbf{M}}) \det(\hat{\kappa}) . \quad (66)$$

On the other hand, comparison of (61) and (63) yields

$$\det(\hat{p}_0 \mathbf{1} - \hat{p}_\rho \hat{\mathbf{L}}^\rho) \det(\hat{\mathbf{M}}) = \det(\tilde{p}_0 \mathbf{1} - \tilde{p}_\rho \tilde{\mathbf{L}}^\rho) \det(\hat{\mathbf{M}}) \quad (67)$$

and thus, after dividing by $\hat{p}_0^2 = \tilde{p}_0^2$ and using (66),

$$\prod_{A=1}^4 (\hat{p}_0 - \hat{\omega}_A(\vec{\hat{p}})) \det(\tilde{\kappa}) = \prod_{A=1}^4 (\tilde{p}_0 - \tilde{\omega}_A(\vec{\tilde{p}})) \det(\tilde{\kappa}) . \quad (68)$$

This shows that, also with respect to transformations (46) and (47), the \mathcal{G}^{abcd} of (37) transform as a tensor density of weight 1,

$$\hat{\mathcal{G}}^{abcd} \hat{p}_a \hat{p}_b \hat{p}_c \hat{p}_d = \left| \det \left(\frac{\partial \tilde{x}}{\partial \hat{x}} \right) \right| \tilde{\mathcal{G}}^{abcd} \tilde{p}_a \tilde{p}_b \tilde{p}_c \tilde{p}_d , \quad (69)$$

where in this case the determinant on the right-hand side is equal to 1.

VIII. CALCULATING THE ROOTS OF THE CHARACTERISTIC EQUATION

In this section we will discuss the question of how to calculate the roots of the characteristic equation if the constitutive matrix is given.

$\omega_1(\vec{p})$, $\omega_2(\vec{p})$, $\omega_3(\vec{p})$ and $\omega_4(\vec{p})$ are, together with $\omega_5(\vec{p}) = \omega_6(\vec{p}) = 0$, the six eigenvalues of the matrix $p_\rho \mathbf{L}^\rho$. In general they are complex. Now the sum of all eigenvalues is the trace, the sum of the squares of the eigenvalues is the trace of the square, and so on. (This is obvious in the case of a diagonalizable matrix. It is also true in general, as can be seen from the Jordan decomposition theorem.) Thus, if we define

$$L^{\rho_1 \cdots \rho_i} = \text{trace}(\mathbf{L}^{\rho_1} \cdots \mathbf{L}^{\rho_i}) , \quad (70)$$

where round brackets around indices mean symmetrization, we get

$$\begin{aligned} p_\rho L^\rho &= \omega_1(\vec{p}) + \omega_2(\vec{p}) + \omega_3(\vec{p}) + \omega_4(\vec{p}) , \\ p_\rho p_\sigma L^{\rho\sigma} &= \omega_1(\vec{p})^2 + \omega_2(\vec{p})^2 + \omega_3(\vec{p})^2 + \omega_4(\vec{p})^2 , \\ p_\rho p_\sigma p_\tau L^{\rho\sigma\tau} &= \omega_1(\vec{p})^3 + \omega_2(\vec{p})^3 + \omega_3(\vec{p})^3 + \omega_4(\vec{p})^3 , \\ p_\rho p_\sigma p_\tau p_\lambda L^{\rho\sigma\tau\lambda} &= \omega_1(\vec{p})^4 + \omega_2(\vec{p})^4 + \omega_3(\vec{p})^4 + \omega_4(\vec{p})^4 . \end{aligned} \quad (71)$$

If the constitutive matrix is known, \mathbf{L}^ρ can be calculated from (17) and the $L^{\rho_1 \cdots \rho_i}$ can be calculated from (70). Then, the four equations (71) of (maximal) order four determine the four roots $\omega_A(\vec{p})$. In general, solving fourth-order equations leads to rather awkward expressions. For several special cases, however, this method allows to calculate the roots of the characteristic equation in a convenient way, as will be demonstrated below.

The transformation behaviour (51) of the $\omega_A(\vec{p})$ implies the following transformation behaviour of the $L^{\rho_1 \cdots \rho_i}$ under generalised Galilean transformations (44) and (45).

$$\begin{aligned} c \tilde{p}_\rho \tilde{L}^\rho &= p_\rho (L^\rho + 4v^\rho) , \\ c^2 \tilde{p}_\rho \tilde{p}_\sigma \tilde{L}^{\rho\sigma} &= p_\rho p_\sigma (L^{\rho\sigma} + 2L^\rho v^\sigma + 4v^\rho v^\sigma) , \\ c^3 \tilde{p}_\rho \tilde{p}_\sigma \tilde{p}_\tau \tilde{L}^{\rho\sigma\tau} &= p_\rho p_\sigma p_\tau (L^{\rho\sigma\tau} + 3L^{\rho\sigma} v^\tau + 3L^\rho v^\sigma v^\tau + 4v^\rho v^\sigma v^\tau) , \\ c^4 \tilde{p}_\rho \tilde{p}_\sigma \tilde{p}_\tau \tilde{p}_\lambda \tilde{L}^{\rho\sigma\tau\lambda} &= p_\rho p_\sigma p_\tau p_\lambda (L^{\rho\sigma\tau\lambda} + 4L^{\rho\sigma\tau} v^\lambda + 6L^{\rho\sigma} v^\tau v^\lambda + 4L^\rho v^\sigma v^\tau v^\lambda + 4v^\rho v^\sigma v^\tau v^\lambda) . \end{aligned} \quad (72)$$

As an alternative, one can derive (72) by multiplying each side of (53) sufficiently often with itself and then calculating the trace. However, this is much more tedious than using (51).

From the first equation of (72) we read that, by a generalised Galilean transformation with $v^\rho = -\frac{1}{4}L^\rho$, it is always possible to transform L^ρ to $\tilde{L}^\rho = 0$. Also, we read from (72) that with respect to purely spatial transformations ($v^\rho = 0$, $c = 1$) the $L^{\rho_1 \cdots \rho_i}$ behave as contravariant tensor components.

The characteristic equation is uniquely determined by the coefficients $L^\rho, L^{\rho\sigma}, L^{\rho\sigma\tau}, L^{\rho\sigma\tau\lambda}$. As they are totally symmetric, these are $(3 + 6 + 10 + 15) = 34$ independent real numbers. On the other hand, the constitutive matrix has $(6 \times 6) = 36$ independent components. From this observation it follows that different constitutive matrices must yield the same characteristic equation. It is an interesting and important problem to find a necessary and sufficient condition for two constitutive matrices to give the same characteristic equation. This problem is unsolved so far; however, in the next section we will find a partial answer by determining a group action on the set of all constitutive matrices that leaves the characteristic equation invariant.

IX. $SL(2, \mathbb{R})$ ACTION ON CONSTITUTIVE MATRICES

We consider the group $SL(2, \mathbb{R})$ in terms of its natural representation by 6×6 matrices,

$$SL(2, \mathbb{R}) = \left\{ \begin{pmatrix} a \mathbf{1} & b \mathbf{1} \\ c \mathbf{1} & d \mathbf{1} \end{pmatrix} \mid ad - bc = 1 \right\}. \quad (73)$$

This group acts on the set of all real 6×6 matrices by conjugation, i.e., each element $\mathbf{Q} \in SL(2, \mathbb{R})$ maps each constitutive matrix \mathbf{M} onto $\mathbf{M}' = \mathbf{Q}^T \mathbf{M} \mathbf{Q}$. In terms of 3×3 blocks, the group action reads

$$\begin{pmatrix} \boldsymbol{\kappa} & \boldsymbol{\chi} \\ \boldsymbol{\gamma} & \boldsymbol{\nu} \end{pmatrix} \mapsto \begin{pmatrix} \boldsymbol{\kappa}' & \boldsymbol{\chi}' \\ \boldsymbol{\gamma}' & \boldsymbol{\nu}' \end{pmatrix} = \begin{pmatrix} a^2 \boldsymbol{\kappa} + ac(\boldsymbol{\chi} + \boldsymbol{\gamma}) + c^2 \boldsymbol{\nu} & ab \boldsymbol{\kappa} + ad \boldsymbol{\chi} + bc \boldsymbol{\gamma} + cd \boldsymbol{\nu} \\ ab \boldsymbol{\kappa} + bc \boldsymbol{\chi} + ad \boldsymbol{\gamma} + cd \boldsymbol{\nu} & b^2 \boldsymbol{\kappa} + bd(\boldsymbol{\chi} + \boldsymbol{\gamma}) + d^2 \boldsymbol{\nu} \end{pmatrix}. \quad (74)$$

It is obvious that this group action leaves the determinant invariant, $\det(\mathbf{M}') = \det(\mathbf{M})$, that it maps symmetric matrices onto symmetric matrices and that it preserves the difference of the off-diagonal blocks, $\boldsymbol{\chi}' - \boldsymbol{\gamma}' = \boldsymbol{\chi} - \boldsymbol{\gamma}$. The following calculation shows that the group action leaves the characteristic equation invariant.

$$\begin{aligned} \det \left(p_0 \mathbf{1} - \begin{pmatrix} \mathbf{0} & -p_\rho \mathbf{A}^\rho \\ p_\rho \mathbf{A}^\rho & \mathbf{0} \end{pmatrix} \mathbf{Q}^T \mathbf{M} \mathbf{Q} \right) &= \\ \det \left(p_0 \mathbf{1} - \mathbf{Q} \begin{pmatrix} \mathbf{0} & -p_\rho \mathbf{A}^\rho \\ p_\rho \mathbf{A}^\rho & \mathbf{0} \end{pmatrix} \mathbf{Q}^T \mathbf{M} \right) &= \\ \det \left(p_0 \mathbf{1} - \begin{pmatrix} \mathbf{0} & -p_\rho \mathbf{A}^\rho \\ p_\rho \mathbf{A}^\rho & \mathbf{0} \end{pmatrix} \mathbf{M} \right). \end{aligned} \quad (75)$$

In the first step we have used the Sylvester identity according to which $\det(s\mathbf{1} - \mathbf{AB}) = \det(s\mathbf{1} - \mathbf{BA})$ for all scalars s and all $n \times n$ matrices \mathbf{A} and \mathbf{B} . In the second step we have used that

$$\begin{pmatrix} a \mathbf{1} & b \mathbf{1} \\ c \mathbf{1} & a \mathbf{1} \end{pmatrix} \begin{pmatrix} \mathbf{0} & -p_\rho \mathbf{A}^\rho \\ p_\rho \mathbf{A}^\rho & \mathbf{0} \end{pmatrix} \begin{pmatrix} a \mathbf{1} & c \mathbf{1} \\ b \mathbf{1} & a \mathbf{1} \end{pmatrix} = \begin{pmatrix} \mathbf{0} & -p_\rho \mathbf{A}^\rho \\ p_\rho \mathbf{A}^\rho & \mathbf{0} \end{pmatrix} \quad (76)$$

if $ad - bc = 1$, as can be quickly verified by multiplying out the the left-hand side.

As $SL(2, \mathbb{R})$ is 3-dimensional, the orbits of the group action must be of dimension ≤ 3 . To calculate the dimension of the orbits we have to differentiate the group action. A quick calculation shows that the tangent space to the orbit through the matrix \mathbf{M} with 3×3 blocks according to (15) is spanned by the three matrices

$$\mathbf{E}_1 = \begin{pmatrix} \boldsymbol{\kappa} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\nu} \end{pmatrix}, \quad \mathbf{E}_2 = \begin{pmatrix} \boldsymbol{\gamma} + \boldsymbol{\chi} & \boldsymbol{\nu} \\ \boldsymbol{\nu} & \mathbf{0} \end{pmatrix}, \quad \mathbf{E}_3 = \begin{pmatrix} \mathbf{0} & \boldsymbol{\kappa} \\ \boldsymbol{\kappa} & \boldsymbol{\gamma} + \boldsymbol{\chi} \end{pmatrix}. \quad (77)$$

$\mathbf{E}_1, \mathbf{E}_2$ and \mathbf{E}_3 are linearly independent unless one of the three matrices $\boldsymbol{\kappa}, \boldsymbol{\nu}$ and $\boldsymbol{\gamma} + \boldsymbol{\chi}$ is zero and the other two are linearly dependent. This demonstrates that the group action foliates a dense and open subset of the set of all real 6×6 matrices into three-dimensional orbits.

The $SL(2, \mathbb{R})$ transformations on constitutive matrices contain two interesting special examples. The first is the one-parameter family of transformations with $a = b = 0$ and $b = c^{-1}$ which corresponds to the so-called *reciprocity transformations*. By definition, a reciprocity transformation is a transformation $F_{ab} \mapsto -\frac{1}{\zeta} H_{ab}$, $H_{ab} \mapsto \zeta F_{ab}$ of field strength and excitation, where ζ is a nowhere vanishing pseudoscalar field. A reciprocity transformation changes

$$\vec{E} \mapsto \frac{1}{\zeta} \vec{H}, \quad \vec{B} \mapsto -\frac{1}{\zeta} \vec{D}, \quad \vec{H} \mapsto -\zeta \vec{E}, \quad \vec{D} \mapsto \zeta \vec{B}, \quad (78)$$

and, thus, the constitutive matrix according to

$$\begin{pmatrix} \boldsymbol{\kappa} & \boldsymbol{\chi} \\ \boldsymbol{\gamma} & \boldsymbol{\nu} \end{pmatrix} \mapsto \begin{pmatrix} \frac{1}{\zeta^2} \boldsymbol{\nu} & -\boldsymbol{\gamma} \\ -\boldsymbol{\chi} & \zeta^2 \boldsymbol{\kappa} \end{pmatrix}. \quad (79)$$

This is precisely the transformation produced by the $SL(2, \mathbb{R})$ element with $a = b = 0$ and $b = c^{-1} = \zeta$. So our result contains as a special case the fact that the characteristic equation is invariant under reciprocity transformations (cf. Hehl and Obukhov [6], pp 273). Note, however, that in the case of a nonlinear constitutive law a reciprocity transformation changes the argument of the constitutive matrix.

The second interesting special case is the one-parameter family of transformations with $a = d = 1$ and $c = 0$. This corresponds to adding an *axion* field, i.e., to a transformation $F_{ab} \mapsto F_{ab}$, $H_{ab} \mapsto H_{ab} + \phi F_{ab}$ with a pseudoscalar field ϕ . This transformation changes

$$\vec{E} \mapsto \vec{E}, \quad \vec{B} \mapsto \vec{B}, \quad \vec{D} \mapsto \vec{D} + \phi \vec{B}, \quad \vec{H} \mapsto \vec{H} - \phi \vec{E}, \quad (80)$$

and, thus, the constitutive matrix according to

$$\begin{pmatrix} \kappa & \chi \\ \gamma & \nu \end{pmatrix} \mapsto \begin{pmatrix} \kappa & \chi + \phi \kappa \\ \gamma + \phi \kappa & \nu + \phi(\gamma + \chi) + \phi^2 \kappa \end{pmatrix}. \quad (81)$$

This is precisely the transformation produced by the $SL(2, \mathbb{R})$ element with $a = d = 1$, $c = 0$ and $b = \phi$. We have thus reproduced the known fact (see Hehl and Obukhov [6], p. 265) that adding an axion field does not affect the characteristic equation.

X. REDUCTION TO 3 DIMENSIONS

On the left-hand side of the characteristic equation (30) we have the determinant of a 6×6 matrix. If the impermeability matrix κ or the impermeability matrix ν is invertible, this may be reduced to the determinant of a 3×3 matrix. We give the derivation for the case that κ is invertible. We may use the decomposition (18) of the constitutive matrix. If we feed this into the characteristic equation (30), and apply Sylvester's formula that $\det(s\mathbf{1} - \mathbf{AB}) = \det(s\mathbf{1} - \mathbf{BA})$ for all scalars s and all $n \times n$ matrices \mathbf{A} and \mathbf{B} , we find

$$0 = \det \left(p_0 \mathbf{1} - \begin{pmatrix} \mathbf{1} & \kappa^{-1} \chi \\ \mathbf{0} & \mathbf{1} \end{pmatrix} \begin{pmatrix} \mathbf{0} & -p_\rho \mathbf{A}^\rho \\ p_\rho \mathbf{A}^\rho & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \gamma \kappa^{-1} & \mathbf{1} \end{pmatrix} \begin{pmatrix} \kappa & \mathbf{0} \\ \mathbf{0} & \nu - \gamma \kappa^{-1} \chi \end{pmatrix} \right) \quad (82)$$

$$= \det \begin{pmatrix} p_0 \mathbf{1} - p_\rho (\kappa^{-1} \chi \mathbf{A}^\rho - \mathbf{A}^\rho \gamma \kappa^{-1}) \kappa & p_\sigma \mathbf{A}^\sigma (\nu - \gamma \kappa^{-1} \chi) \\ -p_\rho \mathbf{A}^\rho \kappa & p_0 \mathbf{1} \end{pmatrix}. \quad (83)$$

We now use the well-known rule (see e.g. [19]) that for any $n \times n$ matrices \mathbf{A} , \mathbf{B} , \mathbf{C} , \mathbf{D}

$$\det \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} = \det(\mathbf{AD} - \mathbf{BC}) \quad \text{if} \quad \mathbf{CD} = \mathbf{DC}. \quad (84)$$

This puts the characteristic equation into the form

$$0 = \det(p_0^2 \kappa^{-1} + p_0 p_\rho (\mathbf{A}^\rho \gamma \kappa^{-1} - \kappa^{-1} \chi \mathbf{A}^\rho) + p_\rho p_\sigma \mathbf{A}^\rho (\nu - \gamma \kappa^{-1} \chi) \mathbf{A}^\sigma). \quad (85)$$

If ν is invertible, an analogous calculation results in

$$0 = \det(p_0^2 \nu^{-1} - p_0 p_\rho (\mathbf{A}^\rho \chi \nu^{-1} - \nu^{-1} \gamma \mathbf{A}^\rho) + p_\rho p_\sigma \mathbf{A}^\rho (\kappa - \chi \nu^{-1} \gamma) \mathbf{A}^\sigma). \quad (86)$$

Note that, by a reciprocity transformation (79), equation (85) transforms into (86) and vice versa. Thus, if both ν and κ are invertible, (85) and (86) are indeed equivalent forms of the characteristic equation.

If the constitutive matrix (15) is invertible, (85) and (86) are equivalent to the form derived by Graglia, Uslenghi and Zich [20], eq. (7) and (8), apart from the fact that they considered only linear constitutive laws. If the cross-terms χ and γ vanish, they reduce to the form of Damaskos, Maffett and Uslenghi [21], eq. (7). An alternative version of the characteristic equation, for linear constitutive laws without cross-terms, was derived and discussed by Itin [22].

XI. INVARIANCE UNDER TIME AND SPACE INVERSION

In general, the characteristic equation is not invariant under time inversion $(p_0, p_1, p_2, p_3) \mapsto (-p_0, p_1, p_2, p_3)$. Similarly, it is not invariant under space inversion $(p_0, p_1, p_2, p_3) \mapsto (p_0, -p_1, -p_2, -p_3)$. However, owing to the homogeneity of the characteristic polynomial, it is invariant under combined time and space inversion, $(p_0, p_1, p_2, p_3) \mapsto (-p_0, -p_1, -p_2, -p_3)$. This implies that the characteristic equation is invariant under time inversion if and only if it is invariant under space inversion. It is easy to see from (30), and even more obvious from (85) or (86), that a sufficient

condition for invariance under time inversion is that the magneto-electric cross-terms χ and γ vanish. However, this is not necessary.

Clearly, the characteristic equation is invariant under time inversion if and only if its roots coincide pairwise up to sign, $\omega_4 = -\omega_1$ and $\omega_3 = -\omega_2$. From (71) we read that this is true if and only if $L^\rho = 0$ and $L^{\rho\sigma\tau} = 0$. In this special case (71) reduces to two second order equations for ω_1^2 and ω_2^2 which can be solved easily. The characteristic equation reads

$$0 = p_0^2 (p_0^2 - \omega_1(\vec{p})^2) (p_0^2 - \omega_2(\vec{p})^2) \quad (87)$$

with

$$\omega_{1/2}(\vec{p})^2 = \frac{1}{4} p_\rho p_\sigma L^{\rho\sigma} \pm \sqrt{\frac{1}{4} p_\rho p_\sigma p_\tau p_\lambda (L^{\rho\sigma\tau\lambda} - \frac{1}{4} L^{\rho\sigma} L^{\tau\lambda})} \quad (88)$$

Thus, the necessary and sufficient condition for hyperbolicity in the time-symmetric case is that the right-hand side of (88) is real and non-negative, i.e. $p_\rho p_\sigma L^{\rho\sigma} \geq 0$ and

$$\frac{1}{2} (p_\rho p_\sigma L^{\rho\sigma})^2 \geq p_\rho p_\sigma p_\tau p_\lambda L^{\rho\sigma\tau\lambda} \geq \frac{1}{4} (p_\rho p_\sigma L^{\rho\sigma})^2 \quad (89)$$

for all \vec{p} in \mathbb{R}^3 . If, in addition, we want to prohibit zero-frequency modes, we have to strengthen the condition $p_\rho p_\sigma L^{\rho\sigma} \geq 0$ to

$$p_\rho p_\sigma L^{\rho\sigma} > 0 \quad \text{for all } \vec{p} \neq \vec{0}. \quad (90)$$

It would be desirable to rewrite (89) and (90) as conditions on the constitutive matrix (15). However, it is hard to see how this can be done in a practicable way.

Equation (87), with (88) and (90), is the general form of the characteristic equation for the case that we have two real double-cones which are mirror-symmetric with respect to time inversion and that zero-frequency modes are prohibited. Each of the two double-cones is the null cone of a *Finsler metric*

$$g_A^{ab} = \frac{1}{2} \frac{\partial^2 (-p_0^2 + \omega_A(\vec{p})^2)}{\partial p_a \partial p_b}, \quad A = 1, 2. \quad (91)$$

More explicitly, the time-time, time-space and space-space components of the two Finsler metrics g_1 and g_2 read

$$g_{1/2}^{00} = -1, \quad g_{1/2}^{0\mu} = 0, \quad g_{1/2}^{\rho\sigma} = \frac{1}{4} L^{\rho\sigma} \pm \frac{\sqrt{\frac{1}{4} p_\rho p_\sigma p_\tau p_\lambda (L^{\rho\sigma\tau\lambda} - \frac{1}{4} L^{\rho\sigma} L^{\tau\lambda})}}{L^{\mu\nu} p_\mu p_\nu} L^{\rho\sigma}. \quad (92)$$

The conditions (89) and (90) guarantee that each of these two Finsler metrics has Lorentzian signature at all $\vec{p} \neq \vec{0}$. For such Finsler metrics a Fermat principle was proven in [23]. Note, however, that in this article the Finsler light cones were assumed to be smooth everywhere (except, of course, at the vertex). This is not the case with the metrics (92). At points where the square-root in (92) has an isolated zero the two light cones form conical singularities. The resulting phenomenon of “conical refraction” was already mentioned in Section VI.

XII. CONDITION OF NON-BIREFRINGENCE

If there are two different real roots $\omega_1(\vec{p}) \geq 0$ and $\omega_2(\vec{p}) \geq 0$, we have birefringence in the forward direction; similarly, if there are two different real roots $\omega_3(\vec{p}) \leq 0$ and $\omega_4(\vec{p}) \leq 0$, we have birefringence in the past direction. In this section we want to investigate the condition for non-birefringence. To that end we consider the case that the four roots ω_1 , ω_2 , ω_3 , and ω_4 pairwise coincide, $\omega_1(\vec{p}) = \omega_2(\vec{p})$ and $\omega_3(\vec{p}) = \omega_4(\vec{p})$ for all $\vec{p} \in \mathbb{R}^3$. For the time being we do not require that the roots are real. With the help of our assumption that the roots pairwise coincide, it is easy to solve the first two equations of (71),

$$\omega_{1/3}(\vec{p}) = \frac{1}{4} p_\rho L^\rho \pm \sqrt{\frac{1}{4} p_\rho p_\sigma (L^{\rho\sigma} - \frac{1}{4} L^\rho L^\sigma)}. \quad (93)$$

By a generalised Galilean transformation we can always transform L^ρ to zero, recall (72), so that (93) simplifies to

$$\omega_{1/3}(\vec{p}) = \pm \sqrt{\frac{1}{4} p_\rho p_\sigma L^{\rho\sigma}}. \quad (94)$$

Note that time-symmetry is then automatically satisfied. As a consequence, the reduced characteristic equation reads

$$(p_0 - \omega_1(\vec{p}))^2 (p_0 - \omega_3(\vec{p}))^2 = (p_0^2 - \frac{1}{4} L^{\rho\sigma} p_\rho p_\sigma)^2 = 0. \quad (95)$$

Thus, the characteristic variety is the null-cone of a quadratic form, i.e., our assumption that the roots pairwise coincide excludes proper Finsler structures. The coefficients $L^{\rho\sigma}$ that determine the quadratic form depend on x and, in the case of a non-linear constitutive law, also on $\vec{D}(x)$ and $\vec{B}(x)$. This result is true independent of whether or not we require hyperbolicity, i.e., independent of whether or not the roots (94) are real.

By (94), hyperbolicity is satisfied if and only if $L^{\rho\sigma}$ is positive semidefinite. This gives a Lorentzian or a degenerate quadratic form. The degenerate case is excluded if we require that $\omega_A(\vec{p}) \neq 0$ for $A = 1, 2, 3, 4$ and $\vec{p} \neq \vec{0}$. Thus, the condition of non-birefringence necessarily leads to a Lorentzian null cone if we require hyperbolicity and exclude zero-frequency modes.

These findings corroborate earlier results found by Hehl and Lämmerzahl [24] and by Itin [25] for linear constitutive laws. They give a satisfactory answer to the question of what kind of light cones are possible in the case of non-birefringence. However, it would also be desirable to have a condition on the constitutive matrix that is necessary and sufficient for non-birefringence. Such a condition is still to be found.

XIII. THE SYMMETRIC HYPERBOLIC CASE

Recall that the evolution equations are symmetric hyperbolic if and only if there is a matrix \mathbf{S} such that

$$\mathbf{S}^{-1} p_\rho \mathbf{L}^\rho \mathbf{S} = (\mathbf{S}^{-1} p_\rho \mathbf{L}^\rho \mathbf{S})^T \quad (96)$$

for all \vec{p} in \mathbb{R}^3 , where $(\cdot)^T$ denotes transposition. With \mathbf{L}^ρ from (17), (96) takes the form

$$\begin{pmatrix} 0 & -p_\rho \mathbf{A}^\rho \\ p_\rho \mathbf{A}^\rho & \mathbf{0} \end{pmatrix} \mathbf{M} \mathbf{S} \mathbf{S}^T = \mathbf{S} \mathbf{S}^T \mathbf{M}^T \begin{pmatrix} 0 & -p_\rho \mathbf{A}^\rho \\ p_\rho \mathbf{A}^\rho & \mathbf{0} \end{pmatrix}. \quad (97)$$

We want to give a characterization of the symmetric hyperbolic case in terms of the constitutive matrix \mathbf{M} . To that end we use the following result which is based on a simple Schur lemma type argument.

Lemma 5. If \mathbf{U} and \mathbf{V} are real 3×3 matrices such that

$$p_\rho \mathbf{A}^\rho \mathbf{U} = \mathbf{V} p_\rho \mathbf{A}^\rho \quad (98)$$

for all $\vec{p} \in \mathbb{R}^3$, then $\mathbf{U} = \mathbf{V} = c\mathbf{1}$ with some $c \in \mathbb{R}$.

Proof. For any $\vec{p} \neq \vec{0}$ in \mathbb{R}^3 , the matrix $p_\rho \mathbf{A}^\rho$ has a one-dimensional kernel spanned by \vec{p} . Thus, by applying (98) to multiples of \vec{p} we see that \mathbf{U} maps every one-dimensional subspace into itself, hence $\mathbf{U} = c\mathbf{1}$. Then (98) takes the form $(\mathbf{V} - c\mathbf{1}) p_\rho \mathbf{A}^\rho = \mathbf{0}$. If \vec{p} runs over \mathbb{R}^3 , the image of $p_\rho \mathbf{A}^\rho$ runs over \mathbb{R}^3 . Thus, our last equation requires $\mathbf{V} = c\mathbf{1}$ \square

With the help of this lemma, we prove the following proposition.

Proposition 5. The evolution equations are symmetric hyperbolic if and only if the constitutive matrix \mathbf{M} is (i) the zero matrix, (ii) positive definite, or (iii) negative definite.

Proof. By applying Lemma 5 to each of the four 3×3 blocks of (97) we find

$$\mathbf{M} \mathbf{S} \mathbf{S}^T = \begin{pmatrix} b\mathbf{1} & d\mathbf{1} \\ a\mathbf{1} & c\mathbf{1} \end{pmatrix}, \quad (99)$$

$$\mathbf{S} \mathbf{S}^T \mathbf{M}^T = \begin{pmatrix} c\mathbf{1} & -a\mathbf{1} \\ -d\mathbf{1} & b\mathbf{1} \end{pmatrix}, \quad (100)$$

with real numbers a, b, c, d . As the left-hand side of (99) is the transpose of the left-hand side of (100), this can be true only if $a = d = 0$ and $b = c$, i.e. the necessary and sufficient condition for symmetric hyperbolicity is that there exists an invertible matrix \mathbf{S} such that $\mathbf{M} = c(\mathbf{S} \mathbf{S}^T)^{-1}$. This is true (i) with $c = 0$ if and only if $\mathbf{M} = \mathbf{0}$, (ii) with $c > 0$ if and only if \mathbf{M} is positive definite, and (iii) with $c < 0$ if and only if \mathbf{M} is negative definite. \square

Clearly, the case $\mathbf{M} = \mathbf{0}$ yields the characteristic equation $p_0^6 = 0$ and is physically uninteresting. Thus, \mathbf{M} must be positive or negative definite to give symmetric hyperbolic evolution equations. In the case of a linear constitutive law this condition is equivalent to the assumption that the *energy density* $w = \frac{1}{2}(E_\rho D^\rho + \mathcal{H}_\rho B^\rho)$ is positive or negative definite. In particular, the positive (or negative) definiteness of \mathbf{M} requires that $\boldsymbol{\kappa}$ and $\boldsymbol{\nu}$ are positive (or negative) definite, i.e., that we have positive (or negative) definite permittivity $\boldsymbol{\varepsilon} = \boldsymbol{\kappa}^{-1}$ and permeability $\boldsymbol{\mu} = \boldsymbol{\nu}^{-1}$.

Proposition 5 generalises a result that was derived in [26], Section 2.1. There only linear constitutive laws were considered and a Lorentzian metric was presupposed. It was shown that, if magneto-electric cross-terms are absent and permeability and permittivity are positive definite, the evolution equations are symmetric hyperbolic. The above result shows that the definiteness condition is not only sufficient but also necessary and that this result (i) carries over to non-linear local constitutive laws, (ii) can be formulated without reference to a background metric, and (iii) remains true if magneto-electric cross-terms are allowed.

XIV. EXAMPLES

A. Biisotropic media

A medium is called *biisotropic* (at a point x) if there is coordinate system such that each 3×3 block of the constitutive matrix is a scalar multiple of the unit matrix (at x), i.e.

$$\mathbf{M} = \begin{pmatrix} \boldsymbol{\kappa} & \boldsymbol{\chi} \\ \boldsymbol{\gamma} & \boldsymbol{\nu} \end{pmatrix} = \begin{pmatrix} \kappa \mathbf{1} & \chi \mathbf{1} \\ \gamma \mathbf{1} & \nu \mathbf{1} \end{pmatrix} \quad (101)$$

with scalars $\kappa, \nu, \chi, \gamma$. If this is true with $\chi = \gamma = 0$, the medium is called *isotropic*. We want to find a necessary and sufficient condition for a biisotropic medium to yield hyperbolic evolution equations.

With (101), the characteristic equation becomes

$$\det(p_0^2 \mathbf{1} - p_0 p_\rho \mathbf{A}^\rho (\chi - \gamma) + p_\rho p_\sigma \mathbf{A}^\rho \mathbf{A}^\sigma (\kappa \nu - \chi \gamma)) = 0 \quad (102)$$

as can be read from (85) or (86). The determinant can easily be calculated, resulting in

$$p_0^2 \left((p_0^2 - |\vec{p}|^2 (\kappa \nu - \chi \gamma))^2 + p_0^2 |\vec{p}|^2 (\chi - \gamma)^2 \right) = 0. \quad (103)$$

Hyperbolicity requires that all roots of this equation are real, which is obviously the case if and only if $\chi = \gamma$ and $\kappa \nu - \chi \gamma \geq 0$. In the last inequality we replace the \geq sign by a $>$ sign, as the $=$ sign only gives zero-frequency modes. With this unphysical zero-frequency case omitted, we can thus say that a biisotropic medium yields hyperbolic evolution equations if and only if

$$\chi = \gamma \quad \text{and} \quad \det(\mathbf{M}) > 0. \quad (104)$$

In this case the reduced characteristic equation has two roots of multiplicity 2,

$$\omega_1(\vec{p}) = \omega_2(\vec{p}) = -\omega_3(\vec{p}) = -\omega_4(\vec{p}) = |\vec{p}| \sqrt{\det(\mathbf{M})}. \quad (105)$$

There is no birefringence, i.e., we have a unique future light cone and a unique past light cone. This double-cone is the null cone of the Lorentzian metric

$$g^{00} = -1 \quad g^{0\mu} = 0 \quad g^{\mu\nu} = \sqrt{\det(\mathbf{M})} \delta^{\mu\nu} \quad (106)$$

and, of course, also of any metric that is conformal to this one.

Note that (104) is equivalent to the requirement that the matrix (101) is (positive or negative) definite. From Section XIII we know that then the evolution equations are symmetric hyperbolic. Hence, for a biisotropic medium the condition of hyperbolicity is equivalent to the condition of symmetric hyperbolicity. For an isotropic medium (104) reduces to the condition that $\kappa = \varepsilon^{-1}$ and $\nu = \mu^{-1}$ must have the same sign.

B. Born-Infeld theory

Born-Infeld theory was introduced by Born and Infeld in 1934 [27]. The motivation was to modify standard vacuum electrodynamics in such a way that the field energy in a small ball around a point charge is finite. This was achieved

by assuming a non-linear vacuum constitutive law of the form $H_{ab} = \partial L / \partial F_{ab}$, where L is the Born-Infeld Lagrangian $L = -\sqrt{b^4 + b^2 F_{ab} F^{ab} - *F_{ab} F^{ab}}$. Here one assumes that, as in standard vacuum electrodynamics, a spacetime metric of Lorentzian signature is given: the star is the Hodge operator defined by the spacetime metric and latin indices are raised and lowered with the spacetime metric. b is a constant of nature, called the “absolute field” by Born and Infeld.

As we allowed for non-linear constitutive laws throughout, Born-Infeld theory fits perfectly well into the general scheme considered in this paper. In the following we will apply the results of the preceding sections to Born-Infeld theory, thereby deriving the structure of the characteristic variety (i.e., of the light cones) in the Born-Infeld theory and establishing the result that the Born-Infeld theory admits a well-posed initial-value problem. None of these results is new. (The light cones of the Born-Infeld theory were determined, e.g by Boillat [28]; a proof that the Born-Infeld initial-value problem is well-posed can be found e.g. in Serre [29].) However, the derivations given here are quite different from the ones available in the literature and illustrate the general results given above.

Our first goal is to demonstrate that, for initial values given on a hypersurface that is spacelike with respect to the spacetime metric, the Born-Infeld initial-value problem is well-posed. To that end we choose coordinates (x^0, x^1, x^2, x^3) such that the chosen spacelike hypersurface is given by the equation $x^0 = \text{constant}$. In addition, we may assume that, at some particular point on the hypersurface, the coordinates are pseudo-orthonormal with respect to the spacetime metric. This leaves the freedom of orthogonal transformations of the spatial coordinates on the tangent space of the chosen point. Then, at the chosen point, the Lagrangian takes the form $L = -\sqrt{b^4 + b^2(B^\mu B_\mu - E^\mu E_\mu) - (E^\mu B_\mu)^2}$. Here and in the following, greek indices are raised and lowered with the Kronecker delta. At the chosen point, the constitutive law reads

$$D^\mu = \frac{\partial L}{\partial E_\mu} = \frac{b^2 E^\mu + B^\nu E_\nu B^\mu}{\sqrt{b^4 + b^2(B^\mu B_\mu - E^\mu E_\mu) - (E^\mu B_\mu)^2}}, \quad (107)$$

$$\mathcal{H}^\mu = -\frac{\partial L}{\partial B_\mu} = \frac{b^2 B^\mu - B^\nu E_\nu E^\mu}{\sqrt{b^4 + b^2(B^\mu B_\mu - E^\mu E_\mu) - (E^\mu B_\mu)^2}}. \quad (108)$$

Clearly, in the limit $b \rightarrow \infty$ the non-linear equations (107) and (108) tend to the linear standard vacuum constitutive law $D^\mu = E^\mu$ and $\mathcal{H}^\mu = E^\mu$. (When comparing our equations (107) and (108) with the corresponding equations on page 437 in the original Born-Infeld paper [27], note that there is a sign error in the latter.)

Equations (107) and (108) can be solved for E^μ and \mathcal{H}^μ , which demonstrates that our coordinates are admissible in the sense of Definition 1. The resulting equations, which are found after an elementary though rather tedious calculation, read

$$E^\mu = \frac{\partial W}{\partial D_\mu} = \frac{1}{W}((b^2 + B^\nu B_\nu)D^\mu - B^\nu D_\nu B^\mu) \quad (109)$$

$$\mathcal{H}^\mu = \frac{\partial W}{\partial B_\mu} = \frac{1}{W}((b^2 + D^\nu D_\nu)B^\mu - B^\nu D_\nu D^\mu) \quad (110)$$

where

$$W(\vec{D}, \vec{B}) = \sqrt{(b^2 + B^\rho B_\rho)(b^2 + D^\sigma D_\sigma) - (B^\tau D_\tau)^2} \quad (111)$$

is the Legendre transform of $L(\vec{E}, \vec{B})$ with respect to the pair of variables \vec{E}, \vec{D} , i.e., $W = E_\mu D^\mu - L$. Now the 3×3 blocks of the constitutive matrix take the form

$$\begin{aligned} \kappa_{\alpha\beta} &= \frac{\partial^2 W}{\partial D^\alpha \partial D^\beta} \\ &= \frac{(b^2 + B^\sigma B_\sigma)}{W^3} \left(W^2 \delta_{\alpha\beta} - (b^2 + B^\rho B_\rho) D_\alpha D_\beta - (b^2 + D^\tau D_\tau) B_\alpha B_\beta + B^\lambda D_\lambda (B_\alpha D_\beta + B_\beta D_\alpha) \right), \\ \nu_{\alpha\beta} &= \frac{\partial^2 W}{\partial B^\alpha \partial B^\beta} = \frac{(b^2 + D^\tau D_\tau)}{(b^2 + B^\rho B_\rho)} \kappa_{\alpha\beta}, \\ \gamma_{\alpha\beta} = \chi_{\beta\alpha} &= \frac{\partial^2 W}{\partial B^\alpha \partial D^\beta} = \frac{B^\tau D_\tau}{(b^2 + B^\rho B_\rho)} \kappa_{\alpha\beta} + \frac{B_\alpha D_\beta - B_\beta D_\alpha}{W}. \end{aligned} \quad (112)$$

κ is symmetric, so it has three real eigenvalues with orthogonal eigenvectors. As our coordinate system is fixed only up to orthogonal transformations of the spatial coordinates, we may choose the coordinates such that κ is diagonal. The eigenvalues κ_1, κ_2 and κ_3 of κ are

$$\begin{aligned}\kappa_{1/2} &= \frac{b^2(b^2 + B^\sigma B_\sigma)}{W^3} \left(b^2 + \frac{D^\mu D_\mu + B^\nu B_\nu}{2} \pm \sqrt{\frac{(B^\rho B_\rho - D^\sigma D_\sigma)^2}{4} + (D^\tau B_\tau)^2} \right), \\ \kappa_3 &= \frac{(b^2 + B^\sigma B_\sigma)}{W}.\end{aligned}\quad (113)$$

As they are strictly positive, κ is positive definite for all (\vec{D}, \vec{B}) in \mathbb{R}^6 .

Now we make a coordinate transformation $(x^0, x^1, x^2, x^3) \mapsto (\tilde{x}^0, \tilde{x}^1, \tilde{x}^2, \tilde{x}^3)$ that induces at the chosen point a generalised Galilean transformation (44) and (45) with

$$\mathbf{a} = \mathbf{b}^{-1} = \text{diag}(\sqrt{\kappa_1}, \sqrt{\kappa_2}, \sqrt{\kappa_3}), \quad c = \det(\mathbf{a}) = \sqrt{\kappa_1 \kappa_2 \kappa_3}, \quad v^\sigma = \frac{1}{W} \epsilon^{\sigma\mu\nu} B_\mu D_\nu. \quad (114)$$

By (49), this transforms (112) into

$$\tilde{\kappa}_{\alpha\beta} = \delta_{\alpha\beta}, \quad \tilde{\nu}_{\alpha\beta} = \frac{(b^2 + D^\tau D_\tau)}{(b^2 + B^\rho B_\rho)} \delta_{\alpha\beta}, \quad \tilde{\gamma}_{\alpha\beta} = \tilde{\chi}_{\beta\alpha} = \frac{B^\tau D_\tau}{(b^2 + B^\rho B_\rho)} \delta_{\alpha\beta}. \quad (115)$$

By (18), the constitutive matrix takes the form

$$\tilde{\mathbf{M}} = \begin{pmatrix} \mathbf{1} & \frac{B^\rho D_\rho}{b^2 + B^\tau B_\tau} \mathbf{1} \\ \mathbf{0} & \mathbf{1} \end{pmatrix}^T \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \frac{W^2}{(b^2 + B^\lambda B_\lambda)^2} \mathbf{1} \end{pmatrix} \begin{pmatrix} \mathbf{1} & \frac{B^\rho D_\rho}{b^2 + B^\tau B_\tau} \mathbf{1} \\ \mathbf{0} & \mathbf{1} \end{pmatrix}. \quad (116)$$

As this matrix is obviously positive definite, Proposition 5 proves that, in the twiddled coordinates, the evolution equations are symmetric hyperbolic, so the initial-value problem is, indeed, well-posed.

Note that the coordinate transformation was necessary for achieving our goal. The Galilean boost with v^σ had the effect of killing the antisymmetric part of $\gamma = \chi^T$. In contrast to $\tilde{\mathbf{M}}$, the original constitutive matrix \mathbf{M} was not positive definite for all values of (\vec{D}, \vec{B}) but only for (\vec{D}, \vec{B}) in a certain neighborhood of the origin in \mathbb{R}^6 .

Finally, we want to calculate the Born-Infeld light cones. In the twiddled coordinates, the characteristic equation (85) takes the form

$$0 = \det\left(\tilde{p}_0^2 \mathbf{1} + \frac{W^2}{(b^2 + B^\mu B_\mu)^2} \tilde{p}_\rho \tilde{p}_\sigma \mathbf{A}^\rho \mathbf{A}^\sigma\right) = \tilde{p}_0^2 \left(\tilde{p}_0^2 - \frac{W^2(\tilde{p}_1^2 + \tilde{p}_2^2 + \tilde{p}_3^2)}{(b^2 + B^\mu B_\mu)^2}\right)^2. \quad (117)$$

Thus, there is no birefringence (cf. Boillat [28]); the Born-Infeld theory determines a unique past and a unique future light cone, given by the equation

$$\tilde{G}^{ab} \tilde{p}_a \tilde{p}_b = -\tilde{p}_0^2 + \frac{W^2(\tilde{p}_1^2 + \tilde{p}_2^2 + \tilde{p}_3^2)}{(b^2 + B^\mu B_\mu)^2} = 0. \quad (118)$$

In the original coordinates, in which the light cone of the spacetime metric takes the form $p_0^2 = p_1^2 + p_2^2 + p_3^2$, (118) reads

$$G^{ab} p_a p_b = -\frac{(p_0 - v^\rho p_\rho)^2}{\kappa_1 \kappa_2 \kappa_3} + \frac{W^2}{(b^2 + B^\mu B_\mu)^2} \left(\frac{p_1^2}{\kappa_1} + \frac{p_2^2}{\kappa_2} + \frac{p_3^2}{\kappa_3}\right) = 0. \quad (119)$$

With the v^ρ from (114) and the κ_μ from (113) the equation for the Born-Infeld light cones becomes

$$\frac{(W p_0 - \epsilon^{\rho\mu\nu} p_\rho B_\mu D_\nu)^2}{b^2(b^2 + B^\lambda B_\lambda)} = b^2(p_1^2 + p_2^2 + p_3^2) - \frac{D^\tau D_\tau + B^\kappa B_\kappa}{2} (p_1^2 + p_2^2) - \sqrt{\frac{(B^\rho B_\rho - D^\sigma D_\sigma)^2}{4} + (D^\tau B_\tau)^2} (p_1^2 - p_2^2). \quad (120)$$

Conclusions

In this article we have considered Maxwell's equations with a local constitutive law and we have found some useful results. In particular, we have derived several versions of the characteristic equation and we have worked out a method of how to calculate its roots; moreover, we have conveniently characterised the class of all constitutive laws

that give symmetric hyperbolic evolution equations. However, symmetric hyperbolicity is not necessary for well-posedness of the initial-value problem. If we want to characterise the class of all constitutive laws for which the initial-value problem is well-posed, we need a criterion for strong hyperbolicity. This is an open problem. It would also be desirable to characterise all constitutive laws that give hyperbolic evolution equations. Again, this is an open problem. We were able to characterise the light cones in the case of invariance under temporal or spatial inversions and in the case of birefringence; however, we could not find a condition on the constitutive matrix that is necessary and sufficient for either of these two properties. We have found a certain group of transformations that act on the set of all constitutive matrices and leave the characteristic equation invariant; however, we could not determine the set of all such transformations. So there are a lot of open problems that should be addressed in future work.

Note added in proof: After this paper was submitted the author learned about Schuller, Witte and Wohlfarth [30], Rätzel, Rivera and Schuller [31], and Favaro and Bergamin [32] where important related results were found.

Acknowledgment

I have profited very much from discussions with Friedrich Hehl, Yuri Obukhov and Yakov Itin on the pre-metric approach to electrodynamics, and from seminars on the subject with Robin Tucker, David Burton, Jonathan Gratus and other colleagues in Lancaster.

-
- [1] F. Kottler, Sitzungsber. Akad. Wien IIa **131**, 119 (1922).
 - [2] É. Cartan, *On manifolds with an affine connection and the theory of general relativity* (Bibliopolis, Napoli, 1986), english translation of the french original from 1923/24
 - [3] D. v. Dantzig, Proc. Cambridge Philos. Soc. **30**, 421 (1934).
 - [4] E. Schroedinger, *Space-time structure* (Cambridge University Press, Cambridge, 1950).
 - [5] E. J. Post, *Formal structure of electromagnetics* (North-Holland, Amsterdam, 1962).
 - [6] F. W. Hehl and Y. Obukhov, *Foundations of classical electrodynamics* (Birkhäuser, Basel, 2003).
 - [7] S. Frittelli, C. N. Kozameh, and E. T. Newman, J. Math. Phys. **36**, 4984 (1995).
 - [8] G. d. Rham, *Differentiable manifolds* (Springer, Berlin, 1984).
 - [9] J. A. Kong, J. Opt. Soc. Amer. **64**, 1304 (1974).
 - [10] J. A. Kong, *Theory of electromagnetic waves* (John Wiley, New York, 1975).
 - [11] T. H. O'Dell, *The electrodynamics of magneto-electric media* (North-Holland, Amsterdam, 1970).
 - [12] F. W. Hehl and Y. Obukhov, Phys. Lett. **A334**, 249 (2005).
 - [13] M. Kline and I. W. Kay, *Electromagnetic theory and geometrical optics* (Interscience, New York, 1965).
 - [14] Y. Itin, J. Phys. A **42**, 475402 (2009).
 - [15] M. Born and E. Wolf, *Principles of optics* (Cambridge University Press, Cambridge, 2002), 7th ed.
 - [16] L. Gårding, J. Math. Mech. **8**, 957 (1959).
 - [17] M. E. Taylor, *Pseudodifferential operators and nonlinear PDE* (Birkhäuser, Boston, 1991).
 - [18] I. V. Lindell, A. H. Sihvola, and K. Suchy, J. Electromagn. Waves Appl. **9**, 887 (1995).
 - [19] J. R. Sylvester, Maths. Gazette **84**, 460 (2000).
 - [20] R. D. Graglia, P. L. E. Uslenghi, and R. E. Zich, IEEE Transactions on Antennas and Propagation **39**, 83 (1991).
 - [21] N. J. Damaskos, A. L. Maffett, and P. L. E. Uslenghi, IEEE Transactions on Antennas and Propagation **30**, 991 (1982).
 - [22] Y. Itin, Phys. Lett. A **374**, 1113 (2010).
 - [23] V. Perlick, Gen. Relativ. Gravit. **38**, 365 (2006).
 - [24] F. W. Hehl and C. Lämmerzahl, Phys. Rev. **D 70**, 105022 (2004).
 - [25] Y. Itin, Phys. Rev. **D 72**, 087502 (2005).
 - [26] V. Perlick, *Ray Optics, Fermat's Principle, and Applications to General Relativity*, vol. m61 of *Lecture Notes in Physics. Monographs* (Springer, Berlin, Germany; New York, U.S.A., 2000).
 - [27] M. Born and L. Infeld, Proc. Roy. Soc. London **A 144**, 425 (1934).
 - [28] G. Boillat, J. Math. Phys. **11**, 941 (1970).
 - [29] D. Serre, Arch. Rat. Mech. Anal. **172**, 309 (2004).
 - [30] F. P. Schuller, C. Witte, and M. N. R. Wohlfarth, Ann. Phys. (NY) **325**, 1853 (2010).
 - [31] D. Rätzel, S. Rivera, and F. P. Schuller, Phys. Rev. **D 83**, 044047 (2011).
 - [32] A. Favaro and L. Bergamin, Ann. Phys. (Berlin) **523**, to appear (2011).